S. ADM
BRARY

DEC 17 1942

Journal of the

AMERICAN STATISTICAL ASSOCIATION

CONTENTS *for* DECEMBER 1942

Cost of Living Indexes in Wartime	*Faith M. Williams, Frances R. Rice and Emil D. Schell*	415
Selective Service's Medical Statistics Program	*Oliver Harold Folk*	425
The Measurement of Capacity Utilization	*Irving H. Siegel*	430
Preliminary Population Estimates Based on Ration Book Applications	*T. J. Woofter, Jr.*	437
Recent Developments in Correlation Technique	*Paul S. Dwyer*	441
Use of the Discriminant Function for More Than Two Groups	*Besse B. Day and Marion M. Sandomire*	461
The New Wheat Condition Figures, Based on Weather Factors, for the Prairie Provinces	*Fred H. Sanderson*	473
A Reply	*C. F. Wilson*	483
Distribution of Income in 1935-36	*Rufus S. Tucker*	489
A Punched Card Technique to Obtain Coefficients of Orthogonal Polynomials	*Lila F. Knudsen*	496
A New Multiplicative Seasonal Index	*Julius Shiskin*	507
Seasonal Factors Determined by Difference from Average of Adjacent Months	*Henry A. Latané*	517
Moving Seasonal Indexes	*Dudley J. Cowden*	523
Errors in Card Punching	*W. Edwards Deming, Benjamin J. Tepping and Leon Geoffrey*	525
Correlations Between Functions of Variables	*G. A. Baker*	537
A Comment on Deming's Classification of Problems of Inference	*Leonard A. Salter, Jr.*	540
Carroll W. Doten, 1871-1942	*Floyd E. Armstrong*	543
Book Reviews		544

VOLUME 37 NUMBER 220

American Statistical Association

Organized November 27, 1839
Incorporated 1841

The American Statistical Association is a scientific and educational organization. Its membership is not confined to professional statisticians but includes economists, business executives, research directors, government officials, university professors, and other persons who are seriously interested in the application of statistical methods to practical problems, in the development of more useful methods, and in the improvement of basic statistical data. Engineers, mathematicians, biologists, actuaries, sociologists, psychologists, and representatives of many other professions are included in the membership of the Association. Information about the Association and membership application forms may be secured from the Secretary, R. L. Funkhouser, American University, Washington, D. C.

Increased restrictions upon civilian travel resulting from the acceleration of war activities have necessitated the cancellation of the Annual Meeting of the American Statistical Association scheduled to be held in Cleveland, Ohio, December 29 to 31, 1942.

Chapters of the Association have been organized in important cities in the United States. Information concerning the meetings of these chapters may be secured from the District Representatives listed on another page.

The Editors welcome the submission of articles and notes for possible publication in the JOURNAL. Manuscripts should be sent to the Editor, Journal of the American Statistical Association, American University, Washington, D. C. Books for review should be sent to the Review Editor, Glenn E. McLaughlin, American University, Washington, D. C. Authors who wish suggestions about the preparation of manuscripts and charts or information about editorial policies should address their inquiries to the Editor.

Subscription rate, $6.00 per annum. Single numbers of volumes 23 to 36 $1.50 per copy postpaid, except Proceedings Supplements to volumes 23 to 30, $2.00 per copy postpaid. Prices for sets and discounts for orders in quantity available on request.

Published Quarterly by the AMERICAN STATISTICAL ASSOCIATION

Publication Office: 450 Ahnaip Street, Menasha, Wisconsin. Editorial Office: American University, Washington, D.C. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, section 538, P. L. & R., authorized March 25, 1936. Entered as second class matter at the post office at Menasha Wisconsin.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

DECEMBER, 1942 · $1.50 PER COPY · $6.00 PER ANNUM · VOL. 37 · NO. 220

CONTENTS

ARTICLES

Cost of Living Indexes in Wartime. . . . . FAITH M. WILLIAMS, FRANCES R. RICE AND EMIL D. SCHELL 415

Selective Service's Medical Statistics Program . OLIVER HAROLD FOLK 425

The Measurement of Capacity Utilization . . . IRVING H. SIEGEL 430

Preliminary Population Estimates Based on Ration Book Applications . . . . . T. J. WOOFTER, JR. 437

Recent Developments in Correlation Technique . . . PAUL S. DWYER 441

Use of the Discriminant Function for More than Two Groups . . . . . BESSE B. DAY AND MARION M. SANDOMIRE 461

The New Wheat Condition Figures, Based on Weather Factors, for the Prairie Provinces . . . . . FRED H. SANDERSON 473

A Reply . . . . . C. F. WILSON 483

Distribution of Income in 1935–36 . . . . . RUFUS S. TUCKER 489

A Punched Card Technique to Obtain Coefficients of Orthogonal Polynomials . . . . . LILA F. KNUDSEN 496

A New Multiplicative Seasonal Index . . . . . JULIUS SHISKIN 507

Seasonal Factors Determined by Difference from Average of Adjacent Months . . . . . HENRY A. LATANÉ 517

Moving Seasonal Indexes . . . . . DUDLEY J. COWDEN 523

Errors in Card Punching. . . . . W. EDWARDS DEMING, BENJAMIN J. TEPPING AND LEON GEOFFREY 525

Correlations Between Functions of Variables . . . . . G. A. BAKER 537

A Comment on Deming's Classification of Problems of Inference . . . . . LEONARD A. SALTER, JR. 540

Carroll W. Doten, 1871–1942 . . . . . FLOYD E. ARMSTRONG 543

BOOK REVIEWS

AGRICULTURAL ECONOMICS, BUREAU OF, see CENSUS, BUREAU OF THE.

BACON, MARVIN A. *Income as an Index of the Fiscal Capacity of Michigan Counties.* Mabel Newcomer 553

BLACK, JOHN D. *Parity, Parity, Parity.* M. K. Bennett 547

BLISS, CHARLES A. *Operating Results of Department and Specialty Stores in 1941.* Lawrence C. Lockley. 555

CENSUS, BUREAU OF THE, AND BUREAU OF AGRICULTURAL ECONOMICS. *Changes in Distribution of Manufacturing Wage Earners, 1899–1939.* Oscar L. Endler 558

Continued on next page

DEAN, JOEL, AND JAMES, R. WARREN. *The Long-run Behavior of Costs in a Chain of Shoe Stores: A Statistical Analysis.* Michael Daly 544

DEAN, JOEL. *The Relation of Cost to Output for a Leather Belt Shop.* With a Memorandum on *Certain Problems in the Empirical Study of Costs,* by C. REINOLD NOYES. Harry P. Hartkemeier 554

DEARDORFF, NEVA R., and FRAENKEL, MARTA. *Hospitals and Hospital Patients in New York City.* Harold F. Dorn 552

FRAENKEL, MARTA, see DEARDORFF, NEVA R.

FREEMAN, H. A. *Industrial Statistics.* Joseph Manuele 559

GARVIN, WILFRED J. *Some Factors Affecting the Supply of Milk and Milk Products in Nova Scotia.* Alan MacLeod 550

GAULT, EDGAR H., and GOODMAN, CHARLES S. *The Control of Customer Returns.* Bishop Brown 548

GOODMAN, CHARLES S., see GAULT, EDGAR H.

HUNT, PEARSON, and KOCH, ALBERT R. *Department Store Financial Requirements. 1920–1940.* Lawrence C. Lockley 555

JAMES, R. WARREN, see DEAN, JOEL.

KING, ARNOLD J., MCCARTY, D. E., and MCPEEK, M. *An Objective Method of Sampling Wheat Fields to Estimate Production and Quality of Wheat.* F. X. Schumacher 549

KOCH, ALBERT R., see HUNT, PEARSON.

MCCARTY, DALE E., see KING, ARNOLD J.

MCPEEK, MILES, see KING, ARNOLD J.

MASS-OBSERVATION (Staff of). *Home Propaganda, Change No. 2.* H. C. Brearley 546

NOYES, C. REINOLD, see DEAN, JOEL.

SPENGLER, JOSEPH H. *French Predecessors of Malthus.* E. P. Hutchinson . 556

WALD, ABRAHAM. *On the Principles of Statistical Inference.* Mark Eudey . 557

PUBLICATIONS RECEIVED 561

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 220 DECEMBER, 1942 *Volume 37*

COST OF LIVING INDEXES IN WARTIME

By Faith M. Williams, Frances R. Rice and Emil D. Schell
U. S. Bureau of Labor Statistics

The cost of living statistician is always in a difficult situation in a period when consumers' buying is changing rapidly. The theory of a cost of living index is that it measures changes through time in the cost of the goods and services customarily purchased by a given group, in a given period, always pricing goods of the same quality. Changes in consumer buying in 1942 are occurring much more rapidly than in the 1920's and 1930's, but many of the problems faced by the statistician at the present time are similar to those met over the last two decades. In wartime, however, it is always more difficult than in peace time for cost of living statisticians to mind their p's and q's.

In the latter half of the 1920's, Mr. Ethelbert Stewart who was then Commissioner of Labor Statistics began to make plans for a new study of consumption patterns on the ground that consumption habits had changed considerably since 1917–19 (the period to which the weights for the Bureau's cost of living index applied). The funds were not made available, however, until 1934 and it was only then that the basic survey was begun which made it possible to take full account in the Bureau of Labor Statistics' cost of living index of the changes in consumption habits since 1919.

It has been explained elsewhere (see Bureau of Labor Statistics' Bulletin 699) that the actual computation of the index did perforce take account of many changes in consumer purchases from 1919 to 1935. A number of the goods included in the index, when it was first calculated, entirely disappeared from the market long before 1935. High button shoes were succeeded by oxfords; pajamas replaced nightshirts; dining room suites replaced dining table, buffet and chairs sold separately. It would have been impossible to proceed with the calculation of the index if many such substitutions had not been made.

After the Bureau's study of the expenditures of wage earners and clerical workers was made in 1934–36, a new set of weights was prepared for each of the city indexes which are included in the Bureau's index for the large cities of the United States. These new weights were

linked with old weights as of March 1935, and a new series computed with new weights from that date.

The number of goods and services covered in the Bureau's new index which had never before been included in the cost of living index provides a rough indication of the changes which occurred in the purchases of wage earners and clerical workers in the United States between the two periods to which the weights for the Bureau's index apply, 1917–19 and 1934–36. The most important of the additions are as follows: Automobiles, automobile repairs, gasoline, motor oil, tires and tubes, automobile licenses and taxes, automobile insurance, electric refrigerators, gas refrigerators, washing machines, vacuum cleaners, radios, electric light bulbs, felt-base floor coverings, living room suites, railroad fare, silk and rayon yard goods, certain fresh fruits and green vegetables. In addition various services were included for the first time: hospital room, surgeon's fees for appendectomy and tonsillectomy, domestic service, beauty shop services, and dry cleaning.

In the last few months, some goods that Americans have customarily purchased for the last 15 or 20 years have suddenly disappeared from the market for the duration; serious shortages have developed in certain staple goods, and some of these like sugar are now being rationed; important quality changes have appeared in other goods. As the production machinery of the country is converted to war purposes, the goods and services produced for consumers are undergoing more drastic changes in the course of a single year than occurred in the entire period from 1919 to 1935—but in the opposite direction. If an index had been computed to show changes in the plane of living of the people of the country, it would have shown a sharp increase in 1941, and would be showing a sharp decline in 1942. But an index of the actual plane of living is quite different from an index of the cost of living.

The drastic changes which have occurred in the supply of consumers' goods available to civilians since December 1940 have made necessary a number of changes in the computation of the Bureau of Labor Statistics' cost of living index. These changes have occurred so rapidly that it has frequently been necessary to make changes without an adequate statistical base. It has not been possible, nor would it have been desirable, as it was when the Bureau produced a new cost of living index as of March 1935, to link in a new bill of goods and services at one date. Adjustments in the index to the new conditions, therefore, have been made currently.

An alternative procedure would have been to hold constant the price of the goods no longer available and to make no change in the weighting of the index, regardless of whether goods are available. If

this procedure had been adopted, it would have been possible to invalidate the cost of living index as a tool for adjusting wages to changes in living costs on the ground that it had lost its connection with reality, that it no longer measured the cost of goods actually available.

Rationing adjustments. Under the first major rationing order which applied to civilians in the United States, new automobiles were withdrawn from the market for an indefinite period. The second order provided that special permission would be required to purchase new tires and tubes and that permission would be granted only in exceptional cases. In January, 1942, accordingly, the Bureau of Labor Statistics adjusted its cost of living index to this situation. New automobiles and new tires were dropped from the index. Used automobiles and used, recapped, and retreaded tires were added, with weights representing the best available estimates of the limited extent to which they were being purchased by wage earners in January and February. The relative importance of automobile repairs and public transportation costs was increased in computing the index. Judging from reports from the Bureau's field agents, a smaller amount of money was being spent by the moderate-income urban families early in 1942 for transportation of all types, than had been spent formerly. Accordingly, the relative importance of transportation costs was reduced from 29 to 23 per cent of the miscellaneous index, and from 8 per cent to about 6 per cent of all living costs. The aggregate cost of the eliminated goods was transferred from the transportation group to the group of unallocated items for which it is assumed prices move with the aggregate of costs for all goods and services priced for the index. The relative importance of the other groups of goods and services included in the miscellaneous index, therefore, was increased slightly.

With further rationing orders, prices of all tires were dropped from the index in April. In May, in the area where gasoline was rationed, automobiles were dropped entirely and drastic cuts made in the weights for gasoline and motor oil. In the areas where gasoline was not rationed, the weight for automobiles was cut in half and for gasoline and motor oil, cut one-third. In November, when gasoline rationing will be universal throughout the country, a further adjustment in transportation weights will take place.

The relative importance of the several transportation costs to moderate-income workers in large cities, as it has changed in the computation of the cost of living index, is shown in Table I.

Disappearance of goods. As striking, perhaps, as the changes in consumer buying resulting from actual rationing orders have been the changes in purchasing habits caused by the disappearance of articles

from the market, as the result of the diversion of raw materials to production of war goods. By June of 1942, many articles of housefurnishings were no longer available for purchase from retail stores. Production of many durable household goods had been stopped earlier

TABLE I

PERCENTAGE DISTRIBUTION OF COST OF TRANSPORTATION TO WAGE EARNERS AND LOWER-SALARIED WORKERS

Item	Relative importance in				
	Base period 1935–39	Dec. 15, 1941	Jan. 15, 1942	March 15, 1942	June 15, 1942
Street car fares	31.5	28.9	39.0	38.5	46.9
Automobiles	26.8	30.3	10.9	11.6	5.7
Gasoline	23.4	22.5	28.1	28.0	22.0
Auto taxes and license fees	3.4	3.4	4.4	6.2	7.5
Automobile insurance	3.0	3.1	3.8	3.9	4.8
Motor oil	2.7	2.5	3.1	3.1	2.4
Bus fares	2.7	3.0	4.2	4.1	5.0
Tires and tubes	2.7	2.6	1.7	—	—
Automobile repairs	2.4	2.5	3.3	3.2	4.0
Railroad fares	1.4	1.2	1.5	1.4	1.7

in the year. By mid-June it was virtually impossible to buy (and, therefore, to price) washing machines, electric and gas refrigerators, stoves, and similar articles in many of the large cities of the country. In one city, for example, almost half (in value) of the articles which make up the housefurnishings group in the index could not be priced.

In computing the index, therefore, the durable goods no longer for sale to civilians were eliminated from the comparison of living costs on June 15 in those cities where they could not be priced. The weight for washing machines, refrigerators, and the other household durables which had virtually disappeared was removed from the housefurnishings index and transferred to the group of unallocated expenditures on the assumption that family money formerly spent for those goods is not used to buy other housefurnishings.

Of course, to some extent the money which would have gone to the purchase of new washing machines may now be going to the purchase of second-hand washing machines. All the available evidence, however, indicates that the supplies of second-hand washing machines and other household durables are rapidly being depleted and it is difficult, if not impossible, to obtain prices of second-hand housefurnishings of comparable quality.

Table II shows the relative importance of each group of items in the computation of the entire cost of living index.

TABLE II

RELATIVE IMPORTANCE OF EACH GROUP OF GOODS AND SERVICES IN COMPUTING CHANGES IN TOTAL LIVING COSTS TO WAGE EARNERS AND CLERICAL WORKERS

Group	Relative importance in				
	Base period 1935–39	Dec. 15, 1941	Jan. 15, 1942	March 15, 1942	June 15, 1942
Food	33.9	34.7	35.2	35.1	35.9
Clothing	10.5	11.0	10.9	11.4	11.4
Rent	18.1	17.6	17.5	17.2	16.8
Fuel, electricity, and ice	6.4	6.1	6.0	5.9	5.8
Housefurnishings	4.2	4.4	4.4	4.5	3.6
Miscellaneous	26.9	26.2	26.0	25.9	26.5
Transportation	7.9	7.7	6.1	6.1	4.9
Medical care	3.8	3.5	3.5	3.4	3.4
Household operation, recreation and personal care	11.0	10.8	10.8	10.6	10.5
Gifts, contributions, and other unallocated items*	4.2	4.2	5.6	5.8	7.7
All Items	100.0	100.0	100.0	100.0	100.0

* Costs for this subgroup are treated as moving with costs for all items priced for the cost of living index.

The relative importance of the various groups of goods and services included in the index changes each time the index is computed because of differences in the rate of price change in the different groups. Since January of this year, the disappearance of certain goods from civilian purchases, and the transfer of the cost weights for these items to the "unallocated" group have resulted in additional changes in the group weights.

Problems of pricing in wartime. At least as difficult of solution as the problems of adjusting the weights of a cost of living index are the problems of pricing in wartime the goods and services which it covers. Since 1935, the Retail Price Division of the Bureau of Labor Statistics has been using quite precise commodity specifications in pricing at retail, in order to be assured that prices at two successive periods refer to the same commodities—the same not only in size and appearance, but in quality as well. This has never been easy, because of the inadequacy of retailers' knowledge of their own merchandise and the almost complete lack of grade labeling. The degree of proficiency which the Retail Price Division's field agents have developed in detecting changes in the quality of the goods they price, however, has assured the users of

the index of a considerable degree of accuracy in the comparison of prices on identical articles from month to month.

Where changes in quality are so concealed that the agents cannot detect them, it has never been possible to make any allowances for quality changes, and will not be possible until programs of quality labeling are developed and put into practice.

Scientific testing of consumers' goods for durability and for efficiency in performing the services for which they are intended is still in its infancy. Performance tests have been developed by the Federal government in order to improve its purchasing procedures, and in order to administer the tariff or acts intended to protect the consumer, or to improve trade practices. Other performance tests have been worked out by industrial engineers working for manufacturers or distributors. In some cases these industry tests have been reviewed by the scientific societies concerned and by the American Standards Association, and have become standard practice; in other cases, as for example the test of heat conductivity for textiles, no test method has received the final approval of any scientific group. As far as the present writers have been able to ascertain, no attempts have been made to prepare measures of efficiency for different types of consumers' goods which would combine several different physical measurements. We have found no statistical comparison of the serviceability, for example, of a specified overcoating made of cotton and reprocessed wool with a specified overcoating made entirely of new wool. The textile chemist can supply test results on resistance to abrasion, breaking-strength, and heat conductivity, although the last-named test is not so well standardized as the others. No one has, however, attempted to work out a formula for weighting the results of these tests into figures which can be used to compare the relative efficiency of the two textiles from the point of view of consumer use.

Today the problem is more acute than ever before. The Bureau's agents must watch not only for quality changes, but also for the changes in sizes and package weights which are reported to have been made since price controls have been imposed. For example, the Bureau's agents noted in 1941 in some stores that a box of corn flakes, priced the same as previously, was marked as containing 6 ounces, where it formerly held 8 ounces. This change appeared in the index of food costs as an increase in the price of a half-pound of corn flakes.

The Bureau follows the policy of continually developing new specifications, as goods formerly purchased are no longer on the market. In pricing for the September, 1942, index, over 100 new specifications

were used, for example, 5 grades of rayon stockings and one grade of cotton anklets. Instructions to the field agents on the use of these specifications, and the specifications for one grade of rayon stockings follows:

> *Rayon hosiery.* Hosiery manufacturers have been consulted as regards these specifications. Cotton reinforcements in the toe, heel, sole and welt are being used when the yarn is available because it is more durable. Viscose yarn has been specified for each item. The use of acetate rayon in hose is still in the experimental stage but will probably become more important in the future.
>
> Hose, rayon, 100 denier, 45 gauge. FABRIC: Viscose rayon, high twist; heel, sole and toe reinforced with cotton, 150 denier rayon welt. CONSTRUCTION AND STYLING: Full fashioned, plain knit welt with run stop, full coursed, not more than 5 flare narrowings.

(Prices are obtained separately for brands widely advertised by manufacturers, and brands not advertised or advertised locally only by distributor or manufacturer.)

Figures on mill shipments of hosiery, retailers' reports, and the observations of the Bureau's field agents indicated that stocks of silk stockings in retail stores were running low. On that account the weight on silk stockings in the Bureau's September index was reduced by two-thirds, and rayon stockings and cotton socks were linked into the index. The weights for rayon stockings and cotton socks were estimated from the latest figures on mill shipments.

The computation of the rent index, again, provides much the same theoretical problems. New dwellings must be linked into the index, when dwellings formerly priced are no longer inhabited, or have deteriorated in quality. However, where decreases have occurred in facilities available, with no change in money rent paid, an increase in rent is recorded.

In the past, all rental quotations obtained by the Bureau for inclusion in the index were obtained from rental agencies. At the present time, however, many property owners are withdrawing rental properties from rental management agencies, so that a sample based on rental management agencies would no longer be representative. In addition, there is some evidence that private owners who handle their own properties are not complying as fully with the ceilings established by the OPA as the real estate agencies who make a regular business of rental management. Beginning with September, 1942, a sample of rent quotations obtained from tenants were collected in 13 of the 34 cities, and after a period of experimentation, were included in the index, in place of the reports from management companies.

Taxes in the index. Excise and sales taxes levied directly upon the price of goods and services sold to consumers have always been included, in the computation of the Bureau's cost of living index, as a part of that cost. Taxes applicable at retail are included by adding the amount of tax to the retail price of each commodity in the index on which such taxes have been imposed. Excise taxes levied at the manufacturers' level are automatically included in the price obtained for inclusion in the index, to the extent that those taxes are passed on to the consumer, since the article is priced in retail stores.

The extent of the price rise due to excise taxes, as distinguished from the extent due to other causes, can be determined, and has been published, on several occasions.[1] This procedure is the same as that followed for the British index. In the July, 1942, issue of the *Ministry of Labour Gazette* (p. 141), there appears the following statement:

> The rise of 45 points since the beginning of September, 1939, is equivalent to about 29 per cent. Of these 45 points, about 4 points represent the effect of the increases, since that date, in the taxes on sugar, tobacco, and cigarettes, and matches; and approximately 6 points are due to increases resulting from the Purchase Tax.

Insofar as taxes on real estate are reflected in higher rental costs, these taxes, too, serve to increase the cost of living, as measured by the Bureau's index. Income taxes, on the other hand, are considered as deductions from income and not as current expenditure.

The amount of the tax to moderate-income families can, of course, be readily computed. In Sweden, this is done, and the change in the amount of the income tax paid by families covered by the cost of living index is published. In addition, the total cost of living index in Sweden, unlike that computed for this country, includes the income tax.

WARTIME ADJUSTMENTS IN THE COST OF LIVING INDEXES OF CANADA, GREAT BRITAIN, AND SWEDEN

The Dominion Bureau of Statistics has, as yet, announced no change with regard to the computation of the Canadian cost of living index, and it may, therefore, be inferred that this index reflects the quality changes in the pre-war plane of living. Few durable goods are priced for inclusion in the Canadian index so that it is unlikely that war time

[1] October 15, 1941, report: BLS #14186 "Changes in Cost of Living in Large Cities."
"The cost of goods purchased by wage earners and lower-salaried workers in large cities increased 1.2 per cent from mid-September to mid-October. Advances in rents and in prices of food, clothing, automobiles, and certain housefurnishings, resulted in an increase of about 1.0 per cent over the month. Excise taxes levied by the Revenue Act of 1941 caused the further increase of about 0.2 per cent."

changes in consumption would have had as much effect on the movement of this index as on that of the United States index.

One change in policy was announced with the publication of the July index regarding direct taxes. Formerly, these taxes had been included in the index but in accordance with the Order-in-Council P. C. 6219, the prices of cigarettes and tobacco used in computing the official index do not include taxes imposed on June 24, 1942, under the Special War Revenue Act. An index comparable with the former index, including the tax, is still computed, together with the new index.

The Ministry of Labour which compiles the official cost of living index for Great Britain, describes this index[2] as representing changes in the cost of the average British workman's plane of living in 1914.

The International Labour Office makes the following comment on this procedure, in computing the British food cost index in wartime.[3]

> It is obvious, therefore, that apart from whatever changes occurred between 1914 and the present war in the composite commodity consumed by the average working-class family, the (British) index is unreliable as an index of the cost of food to the average family in war time. Rationing has considerably altered the kinds and amounts of goods bought by the people. Goods are no longer available in the same quantities as before the war, and the cost of living in war time is the cost of buying a different collection of goods from that bought previously.

The index has been criticized from two points of view: (1) on the ground that the increase in the average household expenditure on many of the items included in the index could not have been so large as the rise in the index, since rationing has reduced the amount of food and clothing available; (2) on the ground that the index excludes a number of miscellaneous foods which have risen in price more than those included in the index. These two criticisms, both of which have some validity, have been regarded by the British government as mutually cancelling out.

No information at present is available to the writers on the procedures used in introducing new types of textiles, clothing, and housefurnishings into the British cost of living index. It is known that 95 per cent of British clothing is now produced according to standard specifications in "utility models," and the index now inevitably follows the cost of some of these models.

[2] The weights are based on studies made in 1904 and 1908. The Ministry of Labour made a study of the expenditures of wage earners and lower-salaried workers in 1937–38 for the purpose of revising these weights, but had not completed the revision by the time England declared war on Germany and the supply of consumers' goods in England as a result of the attack on Poland was necessarily greatly reduced.

[3] *Food Control in Great Britain*, International Labor Office. Montreal, 1942.

A number of adjustments have been made in the cost of living index for Sweden computed by the Social Board,[4] to allow for wartime changes in consumption. Thus, the weights have been reduced by $\frac{1}{3}$ for cheese, $\frac{1}{2}$ for eggs, $\frac{1}{6}$ for flour, bread and cereals, and $\frac{1}{7}$ for sugar. While pork consumption has been reduced by not less than $\frac{2}{3}$, the total meat ration approaches pre-war consumption. The amount of coffee allotted to households has been estimated to fall below usual consumption by only $\frac{1}{4}$ in view of the increased use of coffee substitutes. A corresponding reduction has been made in the weight for coffee in the index.

TABLE III

PER CENT INCREASE IN LIVING COSTS BY GROUPS OF ITEMS FROM AUGUST, 1939, TO AUGUST, 1942, IN CANADA, GREAT BRITAIN AND THE UNITED STATES

	All items	Food	Clothing	Housefurnishings	Rent	Fuel and light	Misc.
Canada	16.8	30.5	20.0	16.7*	7.2	13.6	5.7
Great Britain	29.7	16.8	95.2	†	1.2	33.3	46.7
United States	19.1	34.9	24.7	21.1	3.2	8.9	10.7

* Includes services.
† Included with miscellaneous.

Many factors have been responsible for the disparity in the living cost curves in Great Britain, Canada, and the United States. Differences in the time and degree of active participation in the war, the price control policies adopted, and differences in self-sufficiency are inevitably reflected in the cost of living indexes of the three countries. The situation in Britain was most acute and costs in every major category of family expenditure showed far greater increases than occurred in either of the other two countries until the early part of 1942. At the beginning of this year, the British system of subsidizing production of consumers goods became an effective control and, coupled with lend-lease shipments, was instrumental in bringing about a decline in food costs. Food prices in the United States, however, and those in Canada which are to a large degree dependent on those in our country, were at a higher level in July, 1942, compared with the prewar period, than were those in Great Britain. Costs for the other major groups continued at a relatively higher level than those in America and Canada. Table III shows the rise in each of the three countries by groups from August, 1939, to August, 1942.

[4] *Socialia Meddelanden*, November 10, 1941.

SELECTIVE SERVICE'S MEDICAL STATISTICS PROGRAM

By Oliver Harold Folk
Captain, Infantry, U. S. Army
Selective Service System

The selective service system operated for more than a year before the declaration of war, the first American attempt at peacetime conscription. During this period the primary function of the System was to procure a relatively small number of registrants for a peacetime training program. At that time, summary reports prepared by local boards on classification status and related information seemed to meet the necessities of the System. However, as the international situation developed into a global conflict, the necessity for more detailed and extensive information became apparent.

Registrants not deferred for occupational, dependency, or reasons other than physical or mental, are given a physical examination and, from the beginning of the Selective Service, National Headquarters has received copies of the reports of these physical examinations. These records had not been analyzed by the fall of 1941, at which time a Division of Research and Statistics was organized in National Headquarters and was charged with the responsibility of setting up a centralized system of individual records for all registrants. This system required the establishment of a procedure for receiving certain information relative to each registrant similar to the procedure for receiving a copy of the report of physical examination. This was accomplished so that forms for reporting statistical information could, in general, be made as an additional carbon copy of forms prepared for administrative purposes.

The medical statistics program of the Selective Service System is an integrated part of the entire statistical program, which is based on four essential forms: (1) a list of registrants from each local board showing the name, address, order and serial numbers, race and date of birth; (2) a report of the new and previous classification as soon as each registrant is classified or reclassified; (3) the occupational questionnaire, which shows citizenship, marital status, number of dependents, education, work status, employment class and specialized skills in selected occupations; and (4) the report of physical examination. The list of registrants and the occupational questionnaire are received for all registrants, ages 18 to 65. The classification report is received for only those registrants who are militarily liable, ages 20 to 45, and the

report of physical examination is received for only those registrants militarily liable who are not deferred for occupational, dependency, or reasons other than physical or mental.

Punch cards are prepared for each of the reports and contain an identification number for the registrant, which serves as a basis for matching several records for any one registrant. This identification number consists of the local board number plus the registrant's order number. The local board number is made up of the codes for the corps area, state, county, and local board within the county.

The examining procedure has been altered once and the physical standards have been lowered twice since the beginning of the Selective Service System. Initially, registrants not deferred for occupational, dependency, and other reasons were examined by local boards to determine their acceptability for military service. Those classed as available for general military service were sent to Army induction stations for a final-type physical examination which determined whether or not they were to be inducted for military service. About 28,000 examining physicians and some 10,000 examining dentists assisted the 6,443 local boards in performing the local board physical examination. The physical standards for the examination of Selective Service registrants were used by both local boards and Army induction stations. With the outbreak of war, the need for medical and dental services in the armed forces increased considerably, and on January 1, 1942, local boards were instructed to examine registrants for the purpose of eliminating only those with permanently disqualifying defects that are obvious to the examining physician and do not require instruments for their detection. Those not disqualified by the local board are sent to the induction station. On February 13, 1942, the minimum standards for teeth and eyes were lowered, and effective August 1, 1942, certain types of limited service men, Class 1-B, were acceptable.

Until the change in examination procedure on January 1, 1942, the local board prepared a physical examination report and the Army induction station prepared a separate report for each registrant examined. The analysis of these records, therefore, required the matching of both reports to get the final determination as to physical fitness of the registrant. The new procedure instituted a form in which the findings of the induction station, as well as the local board, were entered on one report. The reports used under the old examining procedure contained a limited amount of information in addition to the medical findings, which consisted of race, occupation, birthplace, birthdate, urban or rural community, medical history, and Selective

Service classification. The new reports contain, in addition to these items, education, years experience at present job, income, employment class, marital status, citizenship, and previous military service.

From the report in current use, much information is available which is needed for administrative uses. The analysis of these reports is divided into two parts: first, a progress reporting procedure whereby the registrant's identification, occupation, marital status, number of dependents, birthdate, race, medical defects, induction station, and examination date are punched currently as received, and summaries made monthly for operational and control purposes by national and state headquarters; and second, intensive studies of samples of these records which will provide information for planning and policy purposes. Current reporting is limited to forms used in the new procedure and samples for intensive study from this group of records will be limited to less than five per cent. However, the tabulation of each report received for registrants examined before January 1, 1942, will not likely be undertaken, and accordingly, a ten per cent sample has been drawn for detailed analysis.

In connection with the preparation of the procedure and instructions for the analysis of the medical reports, a survey of the existing medical nomenclatures currently in use revealed that most of them have been prepared for some specialized purpose, usually morbidity or mortality studies. Selective Service, however, examines a cross-section of the male population between 20 and 45 years of age, an age group which has a relatively low morbidity and mortality rate and, consequently, most of the diseases, defects and impairments found in these men are the hidden abnormal health conditions that are usually not included in morbidity and mortality codes. Also a classification of diseases, defects, and impairments for coding the results of the medical examination of Selective Service registrants must reflect the physical standards by which registrants are examined and the administrative needs of the System.

The medical nomenclature prepared by the Selective Service System is based primarily on the frequency and types of diseases, defects and impairments found at the time of examination of registrants. The specific diagnoses made at the time of examination of 16,000 registrants selected proportionately from each state, at random, were transcribed on to cards, which were then sorted into groups to bring together the defects of each large anatomical section or system of the body. The cards in each of the 27 groups were subdivided into smaller classifications. These detailed classifications were reviewed to eliminate over-

lapping, to clarify titles, to group the ill-defined, and to add new classifications where necessary. Classification titles were grouped and code numbers were assigned so that the code would be flexible enough to care for additional age groups and females, if necessary.

This code of 522 rubrics is used in connection with samples drawn for intensive and detailed studies for planning and policy purposes. An abridged code of 98 classifications is used for complete tabulation and the monthly reports. The 522 classifications are convertible to the 98 classifications. The medical code was tried out on a test sample of 30,000 reports and the results of the test were discussed with specialists for the purpose of remedying such of the defects in the code as were evident.

In the mass processing of these records, an inseparable part of the medical code is the index. The more than 38,000 local board examining physicians and dentists described the defects with considerable diversity of terminology. For consistency in reducing these diagnoses and, in many cases, symptoms, to code numbers, it was necessary to include in the index synonymous terms or expressions actually found on the reports of physical examination. This index consists of 15,000 terms and furnishes a guide for coding more than 95 per cent of the entries being received. Additional terms are being added to this index periodically.

The number of times each registrant has been examined is indicated on the medical record so that the most recent examination report for a registrant can be brought together with previous reports to eliminate duplication. The examining physician, after entering the detailed findings of the examination on the report, also entered a summary of the defects in the order of their seriousness for military service. Each defect recorded is tabulated, with the major defect so designated that it can be tabulated separately. With few exceptions, only one defect is coded for each of the 27 anatomical sections or systems of the body. Further, the coordination operation of defect coding and the tabulation of association tables of defects eliminate coding the cause in one section and the defect in another, or the symptom in one and the diagnosis in another. In addition to the reports received for registrants given a complete physical examination, reports have been received for registrants rejected because of the lack of educational qualifications and for those rejected for obvious physical defects, such as idiocy, epilepsy, feeble-mindedness, blindness, deafness, loss of arms or legs, or confinement in institutions for the insane.

Codes for non-medical items on the punch card have been arranged, wherever possible and feasible, to make these data comparable with existing data of a related nature. The occupational code is based on the "Convertibility List of Occupations and Industry Classifications" prepared in 1940 by the Division of Statistical Standards of the Bureau of the Budget and is convertible to occupational information of the Bureau of the Census and of the United States Employment Service.

While these medical records are being processed and analyzed for operational, policy, and planning purposes, this unique set of data presents a rare opportunity to persons interested in health conditions. Accordingly, from time to time, summaries of these medical reports will be released.

THE MEASUREMENT OF CAPACITY UTILIZATION

By Irving H. Siegel
U. S. Bureau of Labor Statistics

Information on the extent to which existing plant facilities are currently being utilized is indispensable to officials guiding our war production effort and to the many statisticians and economists engaged in the "battle of prediction." The usefulness of much of the available data, however, is qualified by the absence of adequate concepts of capacity and by other fundamental weaknesses which become more evident as maximum output is approached. No doubt, statistics compiled by or for the war agencies will supplement the data available from other sources, but they will surely not be accessible to all who could make use of them. This paper is devoted in part to a discussion of certain theoretical aspects of measurement which are of interest, not only to those compiling new series for capacity and utilization, but also to those who must use what they find. From this theoretical vantage ground, some published indications of capacity utilization are then subjected to examination.

Capacity, as defined in available short-term measures of utilization, may range from some achievable norm to a barely approachable ideal, from a hypothetical output under certain specified conditions to some subjectively determined optimum or maximum.[1] The concept is generally applied to the output of a plant as a whole; the measure of utilization is usually based on the final product of a plant or on the activity of a single department which is supposed to typify the activity of the entire plant. The estimate of capacity is generally fixed for a relatively long period, during which no allowance is made for changes in the volume of equipment or for changes in production methods, operating conditions, and the composition of output.

An important consequence of the methods of measurement is the possibility of attaining a level of capacity utilization in excess of 100 per cent. In such an event, the meaning of an indicator is obscure since the percentage corresponding to true "full" capacity is not given. It would seem to be a desirable alternative to define capacity as a flexible maximum which assumes either actual or other specified conditions not conducing to less than actual production. Such a definition yields a significantly calibrated scale, since 100 would always be the greatest

[1] For descriptions of various measures of capacity utilization, see, for example: E. G. Nourse and Associates, *America's Capacity to Produce* (1934); W. L. Thorp, "The Problem of Overcapacity," *Economic Essays in Honor of Wesley Clair Mitchell* (1935), pp. 479–495; and the "1940 Supplement" to the *Survey of Current Business*.

possible utilization percentage. The capacity estimate would continually be adjusted either for comparability with the actual production conditions in each period or for conformity to new ideal conditions. Such adjustments might be made to reflect changes in amount of plant and equipment; diversion or conversion of equipment from the manufacture of one item to another; changes in yield resulting, say, from improvements in raw materials and production techniques; changes in the character of output; changes in the number of shifts and length of workweek; and changes in the volume of work subcontracted. Clearly, many concepts of capacity, all exceeding actual production, are possible. The meaning of an indicator of capacity utilization thus depends on the explicit definition of capacity selected.

In a certain sense, the development of satisfactory short-term concepts of capacity and capacity utilization depends on the clarification of the short-term concept of production. The conventional measure of production in terms of units of finished product may afford but a crude indication of activity in a short period, particularly if the cycle of fabrication is longer than that period and if the flow of the product through successive manufacturing stages is interrupted by storage or for other reasons is not uniform. If appropriate data were available, it would seem preferable to measure output, not in terms of finished products, but in terms of subproducts of sufficiently short stages in the manufacturing process. The latter method would result in a closer correspondence between activity and output. The number of units of finished product would be regarded, in such a case, as the specific output of the "last" stage of fabrication instead of as the output of the entire plant.

The above concept of production has its analogue in capacity. For short periods, it may be more satisfactory to consider capacity utilization in various departments than in the plant as a whole. A measure of capacity in terms of the number of units of finished product would then be applied, not to the plant as a whole, but to the "last" department. Departmental measures of capacity utilization would be particularly useful in industries where it is desirable, for reasons already given, to measure output in terms of subproducts; where the character of productive activity varies seasonally; where the degree of integration of establishments in an industry is not uniform; and where varying amounts of subproducts are obtained from sources outside an industry. A measure of capacity utilization for an entire plant (or for an industry with establishments of dissimilar structure) could theoretically be derived from departmental production and capacity figures.

For practical purposes, it may be sufficient to measure departmental rates of operation in terms of man-hours, employment, or (similar) machine-hours and to measure plant or industry capacity utilization in terms of man-hours or employment. Thus, one might relate total man-hours actually worked to the maximum that could be worked in the same plant (i.e., in the same work space on the same installed equipment) under actual conditions (e.g., during the same number of shifts of comparable duration, etc.) or under superior conditions (e.g., in the same plant operated on a three-shift basis with experienced workers, etc.). Such measures correspond closely to activity and have the advantage or disadvantage of not being influenced by variations in productivity at different levels of operation. They would be particularly useful for war industries, which subcontract much of their work, have a diversified output, and may have insufficient experience to estimate capacity in physical quantities. Caution must be exercised, however, in the construction and interpretation of indicators based on labor[2] or machine-hours.

In the light of the foregoing discussion, some contemporary indications of capacity utilization will now be considered. No measures are currently available for war industries proper or for subsidiary industries manufacturing complex products (e.g., ordnance or machine tools). Because of the length of the production process and the prevalence of subcontracting, it would appear desirable in such industries to measure utilization departmentally or on a man-hour basis. But most references to operation rates—found, for example, in glowing accounts of local successes in the battle of production—involve finished goods. The concepts of capacity and production are generally loose. An increase in finished products accomplished through subcontracting is treated like an increase in production within the industry; and operation at merely 100 per cent of capacity would seem to be a disappointing performance.

The nebulous state of information on capacity utilization in war industries is evident from a few quotations from recent reports. One observer tells us, for example, that a certain company "estimates that it is turning out machines at the rate of five times its estimated capacity—but it isn't sure." Why? Because, the company says, "no one knows what his capacity is until he is doing everything possible with all the

[2] In view of the current labor shortages, for example, man-hours of different kinds and grades applied to different tasks seem less homogeneous than ever. It is important to distinguish between capacity man-hours for an activity pattern corresponding to actual production and capacity man-hours for an ideal pattern. If labor capacity were being estimated for three-shift operation, it would be a mistake to assume three times the capacity of the first shift. (See "Utilization of Plant Facilities Under National Defense Program," *Monthly Labor Review*, November 1941, pp. 1140–1147, and "Working Hours in War Production Plants, February 1942," *Monthly Labor Review*, May 1942, pp. 1061–1065.)

means at his disposal."[3] The same reporter cites a government armor plant which is now producing "far beyond its rated capacity."[4] In May, newspapers reported "capacity output steadily topped" in Canada's war plants. These plants, according to a digest of the Munitions Minister's remarks to Commons, were producing "above maximum rate." Furthermore, "explosive plants were producing as much as 75 per cent in excess of estimate and one gun plant was turning out two and a half times as many guns as it was thought it could produce."[5] The general manager of a plane company recently confided to newspapermen that "production is only 110 per cent of capacity: 'You see,' he added grinning, 'there is a WPB regulation that says we can't make more than 110 per cent of our allotted monthly quota of planes. If I were to tell you that we are producing more than 110 per cent, the WPB—while I don't guess it really would put me in jail—might not like it.'"[6]

Of the industries for which statistical series are available, steel is by far the most important. In war or peace, the fluctuations in the series for this industry are closely chronicled everywhere. The American Iron and Steel Institute measures of capacity utilization will, therefore, be considered in some detail. Very properly, the series, which are available weekly and monthly, are presented for separate "departments"—i.e., for major types of ingots and rolled products.

Since late in 1940, the measure of capacity utilization for all steel ingots has hovered near the 100 per cent level. Almost every rise or fall is deemed worthy of explanation, and it seems to be assumed that 100 is the true limit. But 100 is not a maximum, as is clear from the methods of measurement, from the achievements of individual companies, and from the records of the component varieties of steel. Consequently, even with current facilities, the industry might be expected to surpass 100 and continue to rise to a truer maximum.

The possibility of exceeding 100 per cent is not only implicit in the definition of "rated capacity" but is also conceded by the steel industry. "Rated capacity" represents "maximum output obtainable in normal operation, as based on tonnages actually obtained"; in its derivation, allowance is made for normal periodic interruptions.[7] Recognizing the "conservative" nature of such an estimate, the industry regards as feasible an annual output of 102.5 per cent of capacity.[8] But, if physical

[3] Paul McCrea, "The Machines Behind the Guns," *Nation's Business*, February 1942, p. 19.
[4] *Ibid.*, p. 27.
[5] *New York Times*, May 15, 1942.
[6] *New York Times*, May 19, 1942.
[7] *Steel Facts*, March 1940, p. 5.
[8] *Steel Facts*, February 1941, p. 3; May 1941, p. 2.

facilities alone were considered, this figure would seem to be an underestimate since individual companies have done much better for rather long periods. For example, a Midwest steel company exceeded its "theoretical capacity" by 19 per cent in 1941.[9] During the first four months of 1941, two-thirds of the 35 reporting companies equaled or exceeded their capacity estimates. Some of the plants also operated at comparable levels during many weeks of 1940. One company's rate was 105 per cent for 28 weeks of 1940; "and then, despite an increase in rated capacity at the year-end, continued to operate at an average of 106 per cent of capacity during 14 of the first 15 weeks of 1941." The 17 plants exceeding capacity surpassed it by 3.3 per cent.[10]

It is a curious fact that, despite the failure of the rate for all steel to exceed 100 per cent, the rates for two of the three major types have surpassed that level several times.[11] For example, the percentage for open-hearth steel, the most important type distinguished by the American Iron and Steel Institute, was 101.9 in March 1941, 100.0 in April, and 100.4 in May. The record of electric steel is more phenomenal. Its rate exceeded 100 in six months of 1941 and will approach, if not duplicate, this feat in 1942. In May 1941, when the rate for all steel was only 98.5, the rates for open-hearth and electric steel were 100.4 and 108.4, respectively. In June 1941, when electric steel reached a peak of 110.8, the average for all steel was only 98.1.

The failure of steel production to be maintained at full rated capacity or higher is too often attributed to physical factors entirely, like the desultory flow of scrap or the shutdown of a furnace for repairs. An important "deterrent" receiving too little attention is purely statistical —namely, the practice of revising capacity estimates at half-year intervals rather than as important changes occur. Recent periodic adjustments have tended to reduce the rates for open-hearth and electric steel and to increase the rate for Bessemer. The reduction, however, has outweighed the rise, forcing down the indicator for all steel in January and July. In reality, the figures for these months are more correct than the percentages for the rest of the year, which tend to be too high. Probably because of the infrequency of revision, however, there is a disposition to accept the inflated figures for all steel as proper indications of operating level. That the overstatement may be substantial is seen, for example, from the magnitude of the adjustment in July 1941; the open-hearth rate declined from 99.4 in June to 94.4 in July, the

[9] "Steel for Victory." February 1942 (a one-page addendum to *Steel Facts*).

[10] *Steel Facts*, May 1941, p. 2.

[11] The discussion in this paragraph and the following one is based on monthly releases of the American Iron and Steel Institute.

electric-steel rate declined from 110.8 to 85.7, and the rate for all steel declined from 98.1 to 93.3. Similarly, the January 1942 revision helped reduce the December 1941 rate for open-hearth steel from 99.0 to 95.4, for electric steel from 101.2 to 96.3, and for all steel from 97.9 to 94.7.

It is often assumed that the percentage of capacity utilization in steelworks is indicative of the rate of operation in fabricating departments. The degree of correspondence, however, depends on a multitude of factors, such as demand, balance of capacity, number of shifts and length of workweek, and inventory policy. The dispersion of operation rates is very wide, as may be seen from statistics for March 1942,[12] when the rate for steel ingots was 98.2. In that month, the rates for various steel products were as follows: standard rails, 52.0; hot-rolled strip, 54.8; drawn wire, 90.7; heavy structural shapes, 91.3; tool-steel bars, 96.0; plates (sheared and universal), 139.5; "all other" (miscellaneous), 190.9; and blackplate, 238.1. It is clear that, in rolling mills as well as in steelworks, it is not unusual nowadays for production to exceed estimated capacity. An interesting illustration of the irrelevance of certain capacity estimates under current operating conditions is afforded by a wide strip mill in a Pittsburgh steel company. Though "theoretically rated" at 70,000 tons a month, this mill accounted for 104,000 tons in March in spite of the variety of products scheduled on it (plates, sheets, strip, and blackplate).[13]

Another important industry represented by a current indicator of utilization is cotton goods—or, more precisely, cotton spinning. This measure, compiled by the U. S. Bureau of the Census, is based on spindle hours.[14] The capacity estimates employed in this measure are now obsolete, for they are based on a 5-day, 80-hour week, whereas yarn mills now operate on more intensive schedules. Thus, it is not surprising to find this series so buoyant since the industry's conversion to war production. In every month of 1941, the rate was well over 100, and, in the first quarter of 1942, the average was about 135. The significance of such rates is impalpable so long as the percentage corresponding to a truer estimate of capacity under contemporary conditions is not known. From the rates themselves, we have no way of anticipating the limit of future changes. It should also be noted that, because of the current lack of balance between carding and spinning equipment, a measure of

[12] American Iron and Steel Institute statistics for steel products are not available for public use after March 1942.

[13] *New York Times*, April 13, 1942. (The quantity 104,000, misprinted in the reference edition, has been verified as correct.)

[14] For current figures, see *Survey of Current Business*. The method of computation is shown in the Census report, *Cotton Production and Distribution: Season of 1941* (Bulletin 178), p. 41.

spindle activity may not even be representative of the entire yarn mill.

By way of summary, four main points of this paper are now repeated. First, available short-term indicators of capacity utilization must be interpreted in the light of the stated or unstated assumptions involved in their construction. Second, it is desirable to extend the concept of capacity from the plant to the department, particularly in war industries where subcontracting is employed to increase the volume of finished output. Third, it may be convenient, especially where production cannot be measured readily on a short-term basis, where subcontracting is general, and where the degree of plant integration is not uniform, to define capacity and capacity utilization in terms of man-hours or employment. Finally, it is desirable, however capacity is defined, to contrive a measure of utilization which cannot exceed 100. Refinements such as these would not only improve our statistical artillery for the present war but would also give us better tools for handling the inevitable problem of "excess capacity" in the post-war period.

PRELIMINARY POPULATION ESTIMATES BASED ON RATION BOOK APPLICATIONS

By T. J. Woofter, Jr.
Director of Research, Federal Security Agency

In May 1942 the estimated population of the United States was about 134,400,000. Roughly 97 per cent[1] of this number filed applications for War Ration Book One (Sugar Ration). If, therefore, it is assumed that a 2.8 per cent undercount in applications was evenly distributed in all areas, then a reasonable approximation to a population count is arrived at by adding 2.8 per cent to the recorded registrations. Such estimates compared with the April 1940 Census yield an idea of the volume and direction of population change, the increases including both excess of births over deaths and migration.

General analysis of the results in the light of other data on changes in employment gives some assurance that the estimates of population obtained by increasing ration applications by 2.8 per cent for states are fairly valid for such large units, but are subject to considerable margin of error in smaller populations.

The principal elements of the population which, theoretically, were not included in the registration and which account for the undercount are:

(a) Members of the armed forces except the small number who were eating in their own domestic quarters;

(b) Institutional populations, such as jails, hospitals, asylums, etc.;

(c) Persons habitually eating at restaurants, although some of these probably did register.

Since these groups were probably larger than the undercount, an undetermined amount of overlap and duplication occurred.

The operation of these factors seems to have been fairly even in large areas. The percentage contribution to the armed forces does not fluctuate widely from area to area. Institutional population is slightly variable from state to state, but there is a wide fluctuation of numbers eating at restaurants.

Other factors which would particularly affect the comparisons of local areas with the 1940 Census are:

(1) Institutional populations are highly concentrated in local areas;

[1] Incomplete application data were used as of September 15, 1942, at which time national total registrations were incomplete by the omission of some large New York and Pennsylvania counties and a few scattered counties in other states. When final figures by states are made available by the Office of Price Administration, they will show a slight variation from this preliminary total.

(2) Temporary laborers would register as of their residence in May 1942 rather than their usual place of abode;

(3) There may in some instances have been lack of exact correspondence in local areas covered by registration and Census enumeration areas.

It is for these reasons that a good many local estimators who have a considerable a priori knowledge of population change have reported mixed results in the use of sugar registration figures as a basis for population estimates. Some of these local estimates have been considered satisfactory and some subject to critical doubts. It must, of course, be noted in this connection that local estimates may at times be as wide of the mark as estimates based upon such registrations.

With this warning as to method, the following observations are submitted as to population change in states and large areas, as estimated by the addition of 2.8 per cent to ration book applications.[2]

According to the estimates by states, 31 states and the District of Columbia increased and 17 decreased from April 1940 to May 1942. The sum of the increase was 3,600,000, whereas the excess of births over deaths in these states was roughly 2,100,000, indicating an approximate interstate migration of about one and a half million people.

When the states are divided into three groups with respect to their relation to the war effort, the results appear as in Table I.

Group 1. The group of 15 states ranking at the top in dollar value of defense contracts and the District of Columbia showed very marked increases, except in Kansas and New York.

In the case of Kansas concentrated industrial activity was offset by continued decrease in the large rural population.

New York entered the war period with the largest actual number and the highest percentage of the population unemployed, and it was able from this reservoir and its natural increase to supply the demands of war industry and export several hundred thousand migrants. In addition, retail and wholesale trade and the non-durable goods industries were not markedly stimulated by war, a fact which contributed to the spectacular loss of about 270,000 in New York City alone.

These states are grouped in the Northeastern Seaboard, around the Great Lakes and on the West Coast. The aggregate increase in the District of Columbia and the 13 states of this group which increased was 2,700,000, or about 4.2 per cent—offset partially as we have noted by decreases.

[2] It should be emphasized that this method adds the members of the armed forces back into their community of origin, yielding a figure as to citizens of the community rather than those in actual residence at the time. To estimate the change in actual residence, the number of members contributed to the armed forces would have to be deducted from the estimates as shown in this article.

TABLE I

POPULATION ENUMERATED APRIL 1940, ESTIMATED MAY 1942, AND INCREASE
(By states)

	Enumerated April 1940	Estimated May 1942	Increase
Group 1. Maximum war activity			
California	6,907,387	7,366,154	458,767
Connecticut	1,709,242	1,817,579	108,337
Illinois	7,897,241	8,211,041	313,800
Indiana	3,427,796	3,582,415	154,619
Kansas	1,801,028	1,766,898	−34,130
Maryland	1,821,028	1,946,969	125,941
Massachusetts	4,316,721	4,390,698	73,977
Michigan	5,256,106	5,709,322	453,216
New York	13,479,142	(Not available)	
New Jersey	4,160,165	4,354,536	194,371
Ohio	6,907,612	7,123,006	215,394
Pennsylvania	9,900,180	(Not available)	
Texas	6,414,824	6,656,272	241,448
Virginia	2,677,773	2,871,025	193,252
Washington	1,736,191	1,799,087	62,896
District of Columbia	663,091	754,000*	90,909
Group 2. Medium war activity			
Alabama	2,832,961	2,984,275	151,314
Arkansas	1,949,387	2,033,377	83,990
Delaware	266,505	282,513	16,008
Florida	1,897,414	2,001,000	103,586
Georgia	3,123,723	3,164,313†	40,590
Louisiana	2,363,880	2,504,730	140,850
Maine	847,226	855,774	8,548
Mississippi	2,183,796	2,200,055	16,259
Missouri	3,784,664	3,847,580	62,916
Nevada	110,247	132,133	21,886
North Carolina	3,571,623	3,601,623†	30,000
Oregon	1,089,684	1,096,805	7,121
Rhode Island	713,346	740,728	27,382
South Carolina	1,899,804	1,958,501	58,697
Tennessee	2,915,841	3,012,092	96,251
Utah	550,310	570,410	20,100
West Virginia	1,901,974	1,915,635	13,661
Wisconsin	3,137,587	3,140,877	3,290
Group 3. Negligible war activity			
Arizona	499,261	477,643	− 21,618
Colorado	1,123,296	1,116,819†	− 6,477
Idaho	524,873	491,191	− 33,682
Iowa	2,538,268	2,517,171	− 21,097
Kentucky	2,845,627	2,837,144	− 8,483
Minnesota	2,792,300	2,754,822	− 37,478
Montana	559,456	539,366	− 20,090
Nebraska	1,315,834	1,280,407	− 35,427
New Hampshire	491,524	487,901	− 1,623
New Mexico	531,818	518,376	− 13,442
North Dakota	641,935	611,622	− 30,313
Oklahoma	2,336,434	2,235,835	−100,599
South Dakota	642,961	605,835	− 37,126
Vermont	359,231	352,181	− 7,050
Wyoming	250,742	239,000	− 11,742

* Applications increased 3 per cent.
† Preliminary.

Group 2. Eighteen states with smaller dollar values of industrial contracts, some of which had troop concentrations which attracted civilian population, and most of which have high rates of natural increases constitute the second group. These states showed some increases, though in most instances not so marked as those in Group 1. The aggregate increase of these states was 900,000, or about 2.6 per cent.

Group 3. Fifteen states with negligible war activity all decreased, the aggregate decrease being 400,000, or about 2.5 per cent.

A high degree of concentration is indicated by the examination of a few metropolitan areas known to have had the maximum industrial activity. (These estimates include whole counties, sometimes several counties surrounding metropolitan cities.)

Over 90 per cent of the increase in the 13 increasing states in Group 1 was in 27 metropolitan areas.

In Illinois, for instance, about 85 per cent of the increase was in the Chicago area. Los Angeles, San Francisco Bay, and San Diego included most of California's increase. More than the Massachusetts state total increase was accounted for in the Boston-Lowell-Lawrence area. More than the total Connecticut increase was in New Haven and Hartford Counties.

This high degree of concentration in a few areas indicates that practically all of the interstate movement and a much larger short-range movement has poured into these areas.

In the second group of states increases were not quite so concentrated, but a fairly large proportion of the migration was into a few industrial centers.

The outstanding conclusions from the foregoing analysis are that in general the large industrial states have become larger and that the trend toward metropolitan concentration has been even more accelerated by the war activity.

RECENT DEVELOPMENTS IN CORRELATION TECHNIQUE

By Paul S. Dwyer
University of Michigan

Every author who has computed correlations and regressions involving a number of variables is aware that extensive computational work is demanded. During recent years certain advances in the computational technique have been introduced which tend to make this calculational work but a fraction of what it was previously for many types of problems. Space forbids an exhaustive review of these advances in this paper though the bibliography at the end should be useful to the student who desires to explore some of them. It is the primary aim of this paper to integrate a number of the most useful of these developments into a series of compact and easy techniques for the numerical solution of the classical problems of multiple and partial correlation and regression.

The development of the modern calculating machine and the adaptation of successive addition to the computation of summed products with machines which are primarily adding machines, such as the Hollerith tabulator, have revolutionalized the computation of the Pearsonian correlation coefficient of the zero order during the past decade. However it is not this topic but rather the computation from the correlation matrix of the quantities used in studying multiple and partial correlation which is the subject of this paper. Each technique outlined builds upon the correlation matrix. Of course many of the techniques are applicable, with slight adjustment, to more general least squares and regression problems but various considerations indicate that an introductory paper of this size should be limited to "correlation" techniques.

A necessary condition for the improvement of calculational technique has been the development of the modern computing machine with automatic division, automatic multiplication and the possibility of evaluation of such expressions as $a-bc-de-fg$ and $(a-bc-de-fg)/h$ as a single operation without copying and resetting. But the chief saving in computational time and energy results not so much from the fact that machines can perform these operations, since we do not gain too much if we use these machines in carrying out techniques which are designed for less refined machines, but in building up new techniques which utilize the mechanical developments to the fullest extent and in finding previously unknown general relationships which can be used in simplifying the calculational work. The introduction of an appropriate

notation, which leads one to the discovery of previously unknown important relationships, for example, is apt to be of much greater value than a mechanical improvement in machines.

The advances which are stressed in this paper are

1. An appropriate notation for the general problem.
2. An abbreviation of the Doolittle solution of the normal equations using machines which simplifies the routine and cuts down the calculational work enormously.
3. The identification of terms in this abbreviated Doolittle solution as correlation constants.
4. Concise methods for calculating the inverse of a matrix.
5. The use of the incomplete inverse as a substitute for the back solution.
6. The identification of terms in the incomplete inverse solution with regression values.

It should be remarked that this paper is really the fifth of a series of papers on modern calculational methods and the basic understanding would be clarified by a reading of papers (28) (29) (35) (36) in which the author has shown the development and theory of topics 1, 2, 4 mentioned above. The reader is also referred to the booklet by Smith (37) which has an aim similar to that of the present paper but which does not use the abbreviated Doolittle solution. The abbreviated Doolittle solution simplifies the presentation, and its understanding, greatly since the illustrations are not cluttered up with a lot of unessential details. Smith refers to the use of the incomplete inverse solution as a substitute for the back solution as the "added column" method and this label seems to be as good as any. The careful student is also advised to consult Deming (14) and Ezekiel (32).

NOTATION

A standard notation for the classical multiple correlation and regression constants is used which involves primary and secondary subscripts with a dot separating them. The secondary subscripts indicate the auxiliary variables involved in the regression while the primary subscripts indicate the specific variables under consideration. Thus if r_{ij} is the correlation between x_i and x_j, $r_{ij \cdot kl}$ is the partial correlation between x_i and x_j with the variables x_k and x_l partialled out and, less precisely, $r_{ij \cdot \ldots}$ represents the partial correlation between x_i and x_j resulting from the elimination of all the other variables of the regression under consideration. The multiple correlation of x_i with the remaining variables is then represented by $r_{i \cdot \ldots}$. The values $\beta_{ij \cdot \ldots}$ are

the standard regression coefficients where x_i is the predicted variable. The standard deviation of the variable (sample) is represented by σ_i and the standard deviation of residuals (sample) by σ_{i_i}....

The main calculational work in estimating the functions of the parent consists in calculating the functions of the sample and it is this aspect of the problem which is stressed in this paper. Estimates in the population may then be obtained with the use of appropriate formulas.

The author has built up a notation for the solution of the Doolittle method which utilizes primary and secondary subscripts in a fashion similar to that of the notation of classical correlation and regression. A complete familiarity with this notation, which is explained under the section devoted to the abbreviated Doolittle solution, is necessary if one is to understand the ideas presented rather compactly in the tables and in the later sections.

THE ABBREVIATED DOOLITTLE SOLUTION

The simplification of the calculational technique is due more to the use of a compact, simple, and streamlined solution of the normal equations than to any other single factor. This solution results from eliminating unnecessary entries from the classical Doolittle solution (35) but it also results from the elimination of unnecessary entries from the rival method of single division (28) which Aitken has called the "method of pivotal condensation" (12). Being the logical sequel of improvements in each of these methods, having its entries directly related to determinantal theory (29), being relatively simple to learn and use, being very compact and requiring a minimum of space, and having its entries immediately interpretable as multiple correlation and regression constants, it is unquestionably the method which should be used by those seeking numerical solutions to the classical problems of multiple and partial correlation.

The normal equations are denoted by

$$\sum a_{ij}x_i = a_{n+1,j} \qquad j = 1, 2, \cdots, n$$

where, in the problem under consideration $a_{ij}=r_{ij}$ and the x_i are the standard regression coefficients. The matrix of coefficients is symmetric since $r_{ij}=r_{ji}$ so that entries below the diagonal need not be recorded in setting up the problem.

The equations for $n=3$ are given symbolically in the first three lines of Table I, General. The first column presents the coefficients of x_1, the second of x_2, the third of x_3, while the fourth presents the constant term. The fifth column is a check column and is obtained by adding the en-

tries in the row. The entries below the main diagonal are enclosed in parentheses.

TABLE I

ABBREVIATED DOOLITTLE SOLUTION

Row	General					Illustration				
1	a_{11}	a_{21}	a_{31}	a_{41}	a_{51}	1.000	.313	.280	.495	2.088
2	(a_{12})	a_{22}	a_{32}	a_{42}	a_{52}	(.313)	1.000	.652	.650	2.615
3	(a_{13})	(a_{23})	a_{33}	a_{43}	a_{53}	(.280)	(.652)	1.000	.803	2.735
4	a_{11}	a_{21}	a_{31}	a_{41}	a_{51}	1.000	.313	.280	.495	2.088
5	1	b_{21}	b_{31}	b_{41}	b_{51}	1.0000	.313	.280	.495	2.088
6		$a_{22\cdot1}$	$a_{32\cdot1}$	$a_{42\cdot1}$	$a_{52\cdot1}$		.902	.564	.495	1.961
7		1	$b_{32\cdot1}$	$b_{42\cdot1}$	$b_{52\cdot1}$		1.000	.625	.549	2.174
8			$a_{33\cdot12}$	$a_{43\cdot12}$	$a_{53\cdot12}$			.569	.355	.924
9			1	$b_{43\cdot12}$	$b_{53\cdot12}$			1.000	.624	1.624
10	$\beta_{31\cdot23}$	$\beta_{32\cdot13}$	$\beta_{33\cdot12}$	$a_{44\cdot123}$		.271	.159	.624	.262	
11	$\beta_{31\cdot23}$	$\beta_{32\cdot13}$	$\beta_{33\cdot12}$			1.271	1.159	1.624		

The first step in the solution consists in writing the equations and the check column in the form indicated. This is done in the illustration of Table I. The correlations used here are those obtained by Carver (27) from a study of body measurements of 1,000 male students entering the University of Michigan. The variables are in order, height, shoulder girth, chest girth, and weight.

The next step consists in writing the values a_{i1} in row 4 so that this row is identical with row 1. These values are then divided by the leading coefficient a_{11} and the results recorded in row 5. The entries in the fifth row are then $b_{i1}=(a_{i1}/a_{11})$. In the correlation problem row 5 is identical with row 4 since $a_{11}=1$. In general the entry in the check column obtained by division is identical, aside from rounding errors, with the sum of the entries to the left.

The next step consists in computing the values

$$a_{i2\cdot1} = a_{i2} - a_{i1}b_{21} \quad \text{(or } a_{i2} - a_{21}b_{i1} \text{ if one prefers)}$$

in recording these entries in row 6, and in computing the values $b_{i2\cdot1}=(a_{i2\cdot1}/a_{22\cdot1})$ by dividing the entries of row 5 by the leading coefficient $a_{22\cdot1}$. This process of computing $a_{i2\cdot1}$ is carried out easily if the first row is covered with a ruler or pencil and the first column is covered similarly so that from the top visible entry a_{i2} one subtracts the product of a_{i1} directly beneath and b_{21}. Thus in the illustration $a_{32\cdot1}=.652-(.280)(.313)=.564$. It is important that the machine be used entirely

in this operation since even a hand calculator is capable of performing an $a-bc$ operation without recording and resetting the intermediate figures. The value $a_{i2\cdot1}$ is used as a check. These entries are then divided by $a_{22\cdot1}$ to get the elements of row 7. Reciprocals or locked divisors may be used depending on the calculating machine available. Also the values of the b's may be found by dividing the a's by $a_{22\cdot1}$ as soon as each $a_{i2\cdot1}$ is available, since machines are capable of performing $(a-bc)/d$ operations to the required number of significant places without resetting. As soon as the b's are determined, the check entry is computed and the sum checked.

The next step is similar. The entries of row 8 are obtained by the use of the formula

$$a_{i3\cdot12} = a_{i3} - a_{i1}b_{31} - a_{i2\cdot1}b_{32\cdot1}$$

and the values of the entries in row 9 by the formula

$$b_{i3\cdot12} = \frac{a_{i3\cdot12}}{a_{33\cdot12}}$$

or from a single operation by

$$b_{i3\cdot12} = \frac{a_{i3} - a_{i1}b_{31} - a_{i2\cdot1}b_{32\cdot1}}{a_{33\cdot12}}.$$

Calculation is facilitated if the first two rows and the first two columns are covered up. In the illustration, for example, one familiar with the meaning of the notation recognizes

$$a_{33\cdot12} = 1.000 - (.280)(.280) - (.564)(.625) = .569$$

$$a_{43\cdot12} = .803 - (.495)(.280) - (.495)(.625) = .355$$

$$b_{43\cdot12} = \frac{.355}{.569} = \frac{.803 - (.495)(.280) - (.495)(.625)}{.569} = .624.$$

This process, continued in the general case until all the predicting variables are eliminated, constitutes an abbreviated form of the forward Doolittle solution. It is very easy to apply when thoroughly understood. Forms or templates may be constructed to assist in the selection of appropriate terms if desired.

An abbreviated form of the back solution is next in order and is presented in row 10. The value $x_3 = \beta_{43\cdot12} =$ the value of $b_{43\cdot12}$ in row 9 is placed in row 10 in the third column. Thus in the illustration the value of $x_3 (= \beta_{43\cdot12})$ is .624. The value of x_2 is then found by solving the equa-

tion of row 7, $x_2+b_{32\cdot1}x_3=b_{42\cdot1}$ so that $x_2=b_{42\cdot1}-b_{32\cdot1}\beta_{43\cdot12}$. Now $\beta_{43\cdot12}$ is entered in row 10 directly under the value of $b_{32\cdot1}$ so that it is not difficult to find the terms and compute the value as a single operation on the machine. Thus in the illustration $\beta_{43\cdot12}=.549-(.625)(.624)=.159$. In a similar way it can be shown from row 5 that

$$\beta_{41\cdot23} = b_{41} - b_{31}\beta_{43\cdot12} - b_{21}\beta_{42\cdot13}$$

so that, in the illustration

$$\beta_{41\cdot23} = .495 - (.280)(.624) - (.313)(.159) = .271.$$

The abbreviated back solution in the more general problem proceeds in exactly the same fashion and requires but a single row.

The back solution can be checked by carrying the check column through the back solution in a similar manner. The entries are shown in row 11 and each is respectively one greater than the corresponding regression coefficient. The real check, however, consists in showing that the calculated coefficients satisfy the original normal equations. The form of the solution is ideally adapted to this type of check since it is easy to pick out the entries and compute

$$(1.000)(.271) + (.313)(.159) + (.280)(.624) = .495, \text{ etc.}$$

Actually some of us feel that the row sum check solution is unnecessary with modern abbreviated methods. In the interest of simplicity it is omitted from the illustrations of later sections.

More detail on the abbreviated Doolittle method cannot be given here since its merits are discussed elsewhere but enough must be given so that the reader understands the notation inasmuch as the contents of later sections are expressed largely in terms of this notation.

It should be noted that a variation of the abbreviated Doolittle method known as the Crout method is being used (30) by some. This method uses columns for the a's, rows for the b's, and so permits a compact solution. The mathematical methods of calculation and the formulas are identical with the methods outlined above but the entries are in different places so there is a somewhat different mechanical technique. An outline of the method, with variations in notation, addition of checks and double diagonal entries introduced by the present author, is presented in Table II.

The solution of the normal equations gives the regression coefficients. One may then use the formula

$$r_{4\cdot123} = \sqrt{\beta_{41\cdot23}r_{41} + \beta_{42\cdot13}r_{42} + \beta_{43\cdot12}r_{43}}$$

TABLE II

SECOND FORM OF ABBREVIATED DOOLITTLE SOLUTION
Variation of Crout Method

General					Illustration				
a_{11}	a_{21}	a_{31}	a_{41}	a_{51}	1.000	.313	.280	.495	2.088
a_{12}	a_{22}	a_{32}	a_{42}	a_{52}	.313	1.000	.652	.650	2.615
a_{13}	a_{23}	a_{33}	a_{43}	a_{53}	.280	.652	1.000	.803	2.735
a_{11}	b_{21}	b_{31}	b_{41}	b_{51}	1.000	.313	.280	.495	2.088
a_{21}	$\frac{1.000}{a_{22\cdot1}}$	$b_{32\cdot1}$	$b_{42\cdot1}$	$b_{52\cdot1}$	.313	$\frac{1.000}{.902}$	.625	.549	2.174
a_{31}	$a_{32\cdot1}$	$\frac{1.000}{a_{33\cdot12}}$	$b_{43\cdot12}$	$b_{53\cdot12}$	.280	.564	$\frac{1.000}{.569}$	.624	1.624
a_{41}	$a_{42\cdot1}$	$a_{43\cdot12}$	$a_{44\cdot123}$		.495	.495	.355	.262	
a_{51}	$a_{52\cdot1}$	$a_{53\cdot12}$			2.088	1.961	.924		
$\beta_{51\cdot23}$	$\beta_{52\cdot13}$	$\beta_{53\cdot12}$			.271	.159	.624		

in obtaining the multiple correlation coefficient. In the illustration above, for example

$$r_{4\cdot123} = \sqrt{(.271)(.495) + (.159)(.650) + (.624)(.803)} = .859.$$

However it is possible, as was pointed out by Horst (1), to obtain the multiple correlation coefficient from the forward solution itself without the necessity of a back solution. For the term $a_{44\cdot123}$ is itself the square of the multiple alienation coefficient as is seen from the relation between the Doolittle solution and determinants (29)

$$r_{4\cdot123} = \sqrt{1 - \frac{\Delta}{\Delta_{44}}} = \sqrt{1 - a_{44\cdot123}}.$$

The value $a_{44\cdot123}$ of the illustration of Table I, $1-(.495)^2-(.495)(.549)-(.355)(.624) = .262$ is inserted in row 10. It is at once possible to obtain the multiple correlation coefficient $r_{4\cdot123} = \sqrt{1-.262} = .859$ without a back solution. The multiple correlation coefficient is also obtainable from the formula

$$r_{4\cdot123} = \sqrt{a_{41}b_{41} + a_{42\cdot1}b_{42\cdot1} + a_{43\cdot12}b_{43\cdot12}}.$$

The standard deviation of residuals is obtainable from the formula

$$\sigma_{44\cdot123} = \sigma_4\sqrt{\frac{\Delta}{\Delta_{44}}} = \sigma_4\sqrt{a_{44\cdot\cdots}}$$

if σ_4, the standard deviation of x_4, is known.

SOLUTION OF RELATED EQUATIONS

One forward solution may be used with different back solutions to solve a large number of systems of regression equations. These regression systems which have a common forward solution may be said to be "related."

Kurtz (9) appears to be the first who saw the application of the Doolittle method to the solution of related equations though he did not use an abbreviated form of the Doolittle solution nor did he incorporate Horst's idea of getting multiple correlations from the forward solution.

The solution of related equations using the abbreviated forward and back Doolittle solution is shown in Table III. The illustration is based on the Carver data from 1,000 students with x_1 = height, x_2 = shoulder girth, x_3 = chest girth, x_4 = waist girth, x_5 = right thigh girth, and x_6 = weight.

The abbreviated forward Doolittle solution is presented in the first 15 rows. The back solutions for different regression systems are shown in rows 16–30. The back solution for the six variable problem in which weight is estimated from the other variables is shown in row 20. The regression equation resulting from the omission of right thigh girth is shown in row 19 and is obtained by covering up the entire column 5 and completing the abbreviated back solution as indicated above. Row 18 is obtained by also omitting waist girth, row 17 also shoulder girth, etc. The sequence of equations 16–20 shows how the standard regression coefficients change with the introduction of an additional predicting variable.

The entries under column 6 are the values of a_{66}...., the square of the alienation coefficient. Rows 21–24 exhibit a corresponding set of regression equations when x_6 is predicted from x_1, x_2, x_3, x_4 and rows 25–27 when x_4 is predicted from x_1, x_2, x_3, etc. The predicted variable is enclosed in each case by vertical lines. The different multiple coefficients are shown in column 7 and the multiple correlation coefficients in column 8. Standard deviation of residuals in the population can then be estimated by appropriate formulas.

THE IDENTIFICATION FORMULAS

It appears that the terms of the abbreviated Doolittle solution may be identifiable in terms of correlation and regression constants and, indeed, this is precisely the case. The identification is secured through the

TABLE III

SOLUTION OF RELATED EQUATIONS

Row								
1	1.000	.313	.280	.182	.166	.495		
2	—	1.000	.652	.554	.615	.650		
3	—	—	1.000	.747	.693	.803		
4	—	—	—	1.000	.774	.804		
5	—	—	—	—	1.000	.812		
6	1.000	.313	.280	.182	.166	.495		
7	1.000	.313	.280	.182	.166	.495		
8		.902	.564	.497	.563	.495		
9		1.000	.625	.551	.624	.549		
10			.569	.385	.295	.355		
11			1.000	.677	.518	.624		
12				.432	.234	.201		
13				1.000	.542	.465		
14					.341	.128		
15					1.000	.375		
							r^2	r
16	.495					.755	.245	.495
17	.323	.549				.483	.517	.719
18	.271	.159	.624			.262	.738	.859
19	.293	.100	.309	.465		.168	.832	.912
20	.310	.013	.252	.262	.375	.120	.880	.938
21	.166				.972		.028	.166
22	−.029	.624			.621		.379	.616
23	−.073	.300	.518		.468		.532	.729
24	−.047	.231	.151	.542	.341		.659	.812
25	.182			.967			.033	.182
26	.010	.551		.693			.307	.554
27	−.048	.128	.677	.432			.568	.754
28	.280		.922				.078	.280
29	.084	.625	.569				.431	.657
30	.313	.902					.098	.313

use of formulas which express the constants in terms of entries of the abbreviated Doolittle solution. Thus such formulas as

$$r_{i\cdot\ldots} = \sqrt{1 - \frac{\Delta}{\Delta_{ii}}} = \sqrt{1 - a_{ii\cdot\ldots}}$$

$$\sigma_{ii\cdot\ldots} = \sigma_i \sqrt{\frac{\Delta}{\Delta_{ii}}} = \sigma_i \sqrt{a_{ii\cdot\ldots}}$$

$$\beta_{ij\cdot\cdot\cdot} = \frac{a_{ij\cdot\cdot\cdot}}{a_{jj\cdot\cdot\cdot}} = b_{ij\cdot\cdot\cdot} \quad \text{if} \quad i > j$$

$$\beta_{ji\cdot\cdot\cdot} = \frac{a_{ji\cdot\cdot\cdot}}{a_{ii\cdot\cdot\cdot}} = \frac{a_{ij\cdot\cdot\cdot}}{a_{ii\cdot\cdot\cdot}} \quad \text{if} \quad j > i$$

$$r_{ij\cdot\cdot\cdot} = \sqrt{\beta_{ij\cdot\cdot\cdot}\beta_{ji\cdot\cdot\cdot}} = \frac{a_{ij\cdot\cdot\cdot}}{\sqrt{a_{ii\cdot\cdot\cdot}a_{jj\cdot\cdot\cdot}}}$$

$$= \sqrt{\frac{a_{ij\cdot\cdot\cdot}b_{ij\cdot\cdot\cdot}}{a_{ii\cdot\cdot\cdot}}} = \sqrt{1 - \frac{a_{ii\cdot\cdot\cdot j,j+1}}{a_{ii\cdot\cdot\cdot j}}}$$

permit the immediate identification of various entries of the forward abbreviated Doolittle solution with correlation constants. This is particularly true if the solution is augmented to record the values of $a_{ii\cdot 1}$, $a_{ii\cdot 12}$, $a_{ii\cdot 123}$, etc. in an adjacent column. This involves no additional work except recording since these quantities are computed in the process of computing $a_{ii\cdot\cdot\cdot}$. The expanded solution is shown in Table IV.

TABLE IV

THE AUGMENTED FORWARD SOLUTION

1.000		.313		.280		.182		.166		.495	
—		1.000		.652		.554		.615		.650	
—		—		1.000		.747		.693		.803	
—		—		—		1.000		.774		.804	
—		—		—		—		1.000		.812	
1.000	1.000	.313	1.000	.280	1.000	.182	1.000	.166	1.000	.495	1.000
1.000		.313		.280		.182		.166		.495	
		.902	.902	.564	.922	.497	.967	.563	.972	.495	.755
		1.000		.625		.551		.624		.549	
				.569	.569	.385	.693	.295	.621	.355	.483
				1.000		.677		.518		.624	
						.432	.432	.234	.468	.201	.262
						1.000		.542		.465	
								.341	.341	.128	.168
								1.000		.375	
										.120	.120

The entries in the added columns are successively a_{ii}, $a_{ii\cdot 1}$, $a_{ii\cdot 12}$, etc. These additional entries are useful in computing the conjugate regression coefficients and the partial correlation coefficients as well as a whole set of related multiple correlation coefficients. For the entries of the added columns are themselves the squares of the multiple alienation

coefficients (identical with those of Table III). Also every $b_{ij}\ldots$ is itself a regression coefficient since $b_{ij}\ldots = \beta_{ij}\ldots$ while every $a_{ij}\ldots$ has its $b_{ij}\ldots$ directly underneath it and its $a_{ii}\ldots$ at its right. Hence every term in the above solution is applicable to one or more of the identification formulas above.

In the illustration consider $a_{65\cdot1234} = .128$ and its auxiliary values $b_{65\cdot1234} = .375$ and $a_{66\cdot1234} = .168$. We have at once by substitution in the identification formulas

$$r_{6\cdot1234} = \sqrt{1 - .168} = .912$$

$$\sigma_{6\cdot1234} = \sigma_6\sqrt{.168} = 7.0 \text{ lbs.} \quad \text{with} \quad \sigma_6 = 17.1 \text{ lbs.}$$

$$\beta_{65\cdot1234} = b_{65\cdot1234} = .375$$

$$\beta_{56\cdot1234} = \frac{a_{65\cdot1234}}{a_{66\cdot1234}} = \frac{.128}{.168} = .767$$

and

$$r_{65\cdot1234} = \sqrt{(.375)(.767)} = \frac{.128}{\sqrt{(.341)(.168)}} = \sqrt{\frac{(.128)(.375)}{.168}}$$

$$= \sqrt{1 - \frac{.120}{.168}} = .536.$$

Every other set of entries can be "identified" likewise.

It becomes apparent that the essential entries of this solution of the normal equations (the abbreviated Doolittle solution) are themselves immediately identifiable in terms of correlation constants and are not merely auxiliary steps in the computation of those constants. It is true that these constants are not always the ones desired, for example, one does not get all the regression coefficients of the back solution of Table III directly from the entries of Table IV, but it is important to recognize that (1) the entries of the solution are themselves identifiable correlation constants and (2) these can be used, in many problems, in computing desired sequences of constants with a minimum of time and effort.

PARTIAL CORRELATIONS WITH THE SAME PRIMARY SUBSCRIPTS

The utility of this identification technique is shown by the solution of partial correlation when 1, 2, 3, 4, etc., variables are successively eliminated. Such a solution is outlined in Table V where as an illustration the correlations between height and weight for the Carver data are shown with shoulder girth, chest girth, waist girth and right thigh girth successively eliminated. The results show that the zero order cor-

TABLE V

COMPUTATION OF A SERIES OF PARTIAL CORRELATIONS HAVING THE SAME PRIMARY SUBSCRIPTS

1	r_{21}	r_{31}	r_{41}	r_{51}	r_{61}					Variable
—	1	r_{32}	r_{42}	r_{52}	r_{62}					
—	—	1	r_{43}	r_{53}	r_{63}					$x_3 =$
—	—	—	1	r_{54}	r_{64}					$x_4 =$
—	—	—	—	1	r_{65}	$a_{11\cdot}$	$a_{22\cdot}$	$a_{66\cdot}$	$r_{56\cdot}$	with
a_{11}	a_{21}	a_{31}	a_{41}	a_{51}	a_{61}	a_{55}	a_{65}	a_{66}	r_{55}	
1	b_{21}	b_{31}	b_{41}	b_{51}	b_{61}					$x_1 =$
	$a_{22\cdot1}$	$a_{32\cdot1}$	$a_{42\cdot1}$	$a_{52\cdot1}$	$a_{62\cdot1}$	$a_{55\cdot1}$	$a_{65\cdot1}$	$a_{66\cdot1}$	$r_{56\cdot1}$	
	1	$b_{32\cdot1}$	$b_{42\cdot1}$	$b_{52\cdot1}$	$b_{62\cdot1}$					$x_2 =$
		$a_{33\cdot12}$	$a_{43\cdot12}$	$a_{53\cdot12}$	$a_{63\cdot12}$	$a_{55\cdot12}$	$a_{65\cdot12}$	$a_{66\cdot12}$	$r_{56\cdot12}$	
		1	$b_{43\cdot12}$	$b_{53\cdot12}$	$b_{63\cdot12}$					$x_3 =$
			$a_{44\cdot123}$	$a_{54\cdot123}$	$a_{64\cdot123}$	$a_{55\cdot123}$	$a_{65\cdot123}$	$a_{66\cdot123}$	$r_{56\cdot123}$	
			1	$b_{54\cdot123}$	$b_{64\cdot123}$					$x_4 =$
				$a_{55\cdot1234}$	$a_{65\cdot1234}$	$a_{55\cdot1234}$	$a_{65\cdot1234}$	$a_{66\cdot1234}$	$r_{56\cdot1234}$	
				$(r_{56\cdot1234})$	$a_{66\cdot1234}$					

Illustration from Carver data

1.000	.652	.554	.615	.313	.650					Variable
—	1.000	.747	.693	.280	.803					
—	—	1.000	.774	.182	.804					x_5 = height
—	—	—	1.000	.166	.812					x_6 = weight
—	—	—	—	1.000	.495	$a_{11\cdot}$	$a_{22\cdot}$	$a_{66\cdot}$	$r_{56\cdot}$	with
1.000	.652	.554	.615	.313	.650	1.000	.495	1.000	.495	
1.000	.652	.554	.615	.313	.650					x_1 = shoulder girth
	.575	.386	.292	.076	.379	.902	.292	.578	.404	
	1.000	.671	.508	.132	.659					x_2 = chest girth
		.434	.237	−.042	.190	.892	.242	.328	.447	
		1.000	.546	−.097	.438					x_3 = waist girth
			.344	−.042	.116	.888	.260	.245	.558	
			1.000	−.122	.337					x_4 = right thigh girth
				.883	.274	.883	.274	.205	.645	
				(.645)	.205					

relation of .5 reduces to .4 and then increases to .65. Partial correlation studies of this sort, almost prohibitive by classical computational methods, are now easily and quickly found by arranging the variables represented by the primary subscripts at the right of the form, carrying

the solution through the elimination of the variables indicated by the secondary subscripts and with the continued use of the formula

$$r_{ji\cdot\cdots} = \frac{a_{ij\cdot\cdots}}{\sqrt{a_{ii\cdot\cdots}a_{jj\cdot\cdots}}}.$$

THE INVERSE OF A MATRIX

The inverse of the matrix is of great value in studying the regression problems. The application to groups of equations in which the predicting variables are the same was introduced by Gauss (14) and has been developed by various authors (4) (12) (13). The inverse of the correlation matrix is of great value, too, in expressing the standard errors of the regression coefficients. During recent years there have been a number of interesting methods developed for finding the inverse of a square symmetric matrix (4) (12) (13) (15) (19) (22) (28) (34) (36) (37). The method emphasized here owes much to the contributions of Aitken (12) although other authors (15) (36) (37) have made contributions. A more formal presentation and justification of the method than can be presented is found in the author's paper on "The Evaluation of Linear Forms" (36).

The method is illustrated in Table VI where the correlation coefficients are those of Table I. The identity matrix is set up on the right and the abbreviated Doolittle solution carried through the elimination of the four variables. The entries of the inverse matrix indicated by C_{ij} are then found by adding the products of a and b terms above. Diagonal

TABLE VI

INVERSE OF A SQUARE SYMMETRIC MATRIX

1.000	.313	.280	.495	1.000	0	0	0
—	1.000	.652	.650	0	1.000	0	0
—	—	1.000	.803	0	0	1.000	0
—	—	—	1.000	0	0	0	1.000
1.000	.313	.280	.495	1.000	0	0	0
1.000	.313	.280	.495	1.000	0	0	0
	.902	.564	.495	− .313	1.000	0	0
	1.000	.625	.549	− .347	1.109	0	0
		.569	.355	− .084	− .625	1.000	0
		1.000	.624	− .147	−1.098	1.757	0
			.262	− .271	− .159	− .624	1.000
			1.000	−1.034	− .607	−2.382	3.817
				1.401	− .091	.498	−1.034
				—	1.892	− .719	− .607
				—	—	3.243	−2.382
				—	—	—	3.817

entries, $i=j$, are found by summing products of terms in the i column while non-diagonal entries, $i \neq j$, are found by multiplying an a from the i-th column and a b from the j-th column and adding. Thus

$$C_{11} = (1.000)(1.000) + (-.313)(-.347) + (-.084)(-.147) + (-.271)(-1.034) = 1.401$$

and

$$C_{12} = C_{21} = (1.000)(0) + (-.313)(1.109) + (-.084)(1.098) + (-.271)(-.607) = -.091 \text{ etc.}$$

Smith (37) has called this method the "Added Column" method.

The adjoint can be found by multiplying each element of the inverse by the value of the determinant which is, in the illustration above, $a_{11}a_{22\cdot1}a_{33\cdot12}a_{44\cdot123} = (1.000)(.902)(.569)(.262)$.

REGRESSION COEFFICIENTS AND THEIR STANDARD ERRORS

The standard errors of the regression coefficients can be computed from the diagonal term of the inverse matrix since

$$\sigma_{\beta ij\ldots} = \sqrt{\frac{a_{ij\ldots}}{n-1-m} C_{jj}}$$

where n is the number of observations and m is the number of independent variables. It appears at once that the diagonal terms only of the inverse solution are necessary and not the complete inverse. These can be written in the same row with the correlation coefficients and the value of $a_{ii\ldots}$.

The use of the incomplete inverse is illustrated in Table VII where the regression coefficients of Table VI and their standard errors are computed. The first four columns and the first ten rows are identical with Table I. The last three columns carry out the inverse solution with the values of C_{11}, C_{22}, C_{33} entered in row ten, the values of

$$\frac{a_{44\cdot123}}{1000-1-3}, \quad \frac{a_{44\cdot123}}{996} C_{11}, \quad \frac{a_{44\cdot123}}{996} C_{22}, \quad \frac{a_{44\cdot123}}{996} C_{33}$$

in row 11, and the standard errors in row 12.

The values of the regression coefficients themselves may be found with the use of the inverse without the necessity of a back solution by multiplying the b's of column 4 by the a's of column 5, 6, 7 in turn.

Thus

$$\beta_{41\cdot 23} = (.495)(1.000) + (.549)(-.313) + (.624)(-.084) = .271,$$
$$\beta_{42\cdot 13} = (.495)(0) + (.549)(1.000) + (.624)(-.625) = .159,$$
$$\beta_{43\cdot 12} = (.495)(0) + (.549)(0) + (.624)(1.000) = .624.$$

TABLE VII

REGRESSION COEFFICIENTS AND THEIR STANDARD ERRORS FROM THE INCOMPLETE INVERSE

x_1	x_2	x_3	x_4			
1.000	.313	.280	.495	1.000	0	0
—	1.000	.652	.650	0	1.000	0
—	—	1.000	.803	0	0	1.000
1.000	.313	.280	.495	1.000	0	0
1.000	.313	.280	.495	1.000	0	0
	.902	.564	.495	−.313	1.000	0
	1.000	.625	.549	−.344	1.109	0
		.569	.355	−.084	−.625	1.000
		1.000	.624	−.147	−1.098	1.757
.271	.159	.624	.262	1.121	1.795	1.757
			.000263	.000295	.000472	.000462
				.017	.0022	.021

MORE IDENTIFICATION FORMULAS

More identification formulas are now available since, as Horst (2) and Waugh (8) have shown, it is possible to compute from the inverse matrix (a) all multiple correlations of order n-1, (b) all partial correlations of order n-2, (c) all regression coefficients of order n-2, (d) the standard errors of regression coefficients of order n-2 with the use of the identification formulas

$$r_{i\cdot\cdots} = \sqrt{1 - \frac{1}{C_{ii}}}$$

$$r_{ij\cdot\cdots} = \frac{-C_{ij}}{\sqrt{C_{ii}C_{jj}}}$$

$$\beta_{ij\cdot\cdots} = -\frac{C_{ij}}{C_{ii}}$$

$$\sigma_{\beta ij\cdot\cdots} = \sqrt{\frac{C_{ii}C_{jj} - C_{ij}^2}{n'C_{ii}}}$$

where n' is the number of degrees of freedom.

More than that it can be shown that the a entries on the right hand side of Table VI are themselves regression coefficients and more than that they are the coefficients needed for a complete regression equation. Thus the coefficients of row 12 of Table VI are the negatives of the coefficients of row 18 of Table I (and row 10 of Table VII) and are the negative coefficients of the $4 \cdot 123$ system. Similarly the coefficients of row 9 of Table VI are the negative coefficients of the $3 \cdot 12$ system, etc. It thus appears that the entries of the inverse solution are also identifiable as correlation constants.

SUCCESSIVE MULTIPLE CORRELATION CONSTANTS FROM THE INCOMPLETE INVERSE

We are now ready for the illustration which shows the real value of many of the advances outlined above. One can build up values of regression coefficients, their standard errors, multiple alienation and correlation coefficients, and the standard deviation of residuals for related groups of equations readily from the incomplete inverse solution.

The illustration used is that of Table III with the incomplete inverse added. The regression coefficients themselves are computed easily by forming the summed successive products of the b's of column 6 and the a's of incomplete inverse. Thus

$$\beta_{61} = .495(1.000) = .495$$
$$\beta_{62 \cdot 1} = .495 + (.549)(-.313) = .323$$
$$\beta_{63 \cdot 12} = .323 + (.624)(-.084) = .271 \quad \text{etc.}$$

This process is most easily carried out and it is only necessary to record each entry as the final term is built up. The values in column 6 are the values $a_{ii} \cdot \ldots$ of Table III. The values $C_{ii \cdot j}$ are the corresponding diagonal elements of the inverse matrices of order 1, 2, 3, 4, 5, respectively. They likewise are easily obtained from the products of "a" and "b" terms above them. Thus

$$C_{11 \cdot 1} = (1.000)(1.000) = 1.000$$
$$C_{11 \cdot 2} = C_{11 \cdot 1} + (-.313)(-.347) = 1.109 \quad \text{etc.}$$

From these values it is possible to calculate the standard errors of the regression coefficients on the left.

The values $r^2_{6} \cdot \ldots$, $r_{6} \cdot \ldots$, and $\sigma_{\epsilon_6} \cdot \ldots$ are given at the right. Corresponding values can be readily computed for all the related regressions of Table III.

TABLE VIII

SUCCESSIVE MULTIPLE CORRELATION CONSTANTS FROM THE INCOMPLETE INVERSE

Illustration from Carver data

1.000	.313	.280	.182	.166	.495	1	0	0	0	0	x_0 = weight		
—	1.000	.652	.554	.615	.650	0	1	0	0	0	n = 1000		
—	—	1.000	.747	.693	.803	0	0	1	0	0	M_0 = 139.3 lbs.		
—	—	—	1.000	.774	.804	0	0	0	1	0	σ_0 = 17.07 lbs.		
—	—	—	—	1.000	.812	0	0	0	0	1			
1.000	.313	.280	.182	.166	.495	1.000	0	0	0	0	x_1 = height		
1.000	.313	.280	.182	.166	.495	1.000	0	0	0	0			
	.902	.564	.497	.563	.495	−.313	1.000	0	0	0	x_2 = shoulder girth		
	1.000	.625	.551	.624	.549	−.347	1.109	0	0	0			
		.569	.385	.295	.355	−.084	−.625	1.000	0	0	x_3 = chest girth		
		1.000	.677	.518	.624	−.147	−1.098	1.757	0	0			
			.432	.234	.201	.047	−.128	−.677	1.000	0	x_4 = waist girth		
			1.000	.542	.465	.109	−.296	−1.567	2.315	0			
				.341	.128	.047	−.231	−.151	−.542	1.000	x_5 = right thigh girth		
				1.000	.375	.138	−.677	−.443	−1.589	2.933			
.495					.755	1.000					.245	.495	14.83 lbs.
.323	.549				.483	1.109	1.109				.517	.719	11.86 lbs.
.271	.159	.624			.262	1.121	1.795	1.757			.738	.859	8.74 lbs.
.293	.099	.309	.465		.168	1.126	1.833	2.818	2.315		.832	.912	7.00 lbs.
.310	.013	.253	.262	.375	.120	1.133	1.990	2.885	3.176	2.933	.880	.938	5.91 lbs.

β_{01}					$a_{00\cdot1}$	$C_{11\cdot1}$					$r_{0}^{2}{}_{\cdot1}$	$r_{0\cdot1}$	$\sigma_{e_{0\cdot1}}$
$\beta_{01\cdot2}$	$\beta_{02\cdot1}$				$a_{00\cdot12}$	$C_{11\cdot2}$	$C_{22\cdot1}$				$r_{0}^{2}{}_{\cdot12}$	$r_{0\cdot12}$	$\sigma_{e_{0\cdot12}}$
$\beta_{01\cdot23}$	$\beta_{02\cdot13}$	$\beta_{03\cdot12}$			$a_{00\cdot123}$	$C_{11\cdot2}$	$C_{22\cdot1}$	$C_{33\cdot1}$			$r_{0}^{2}{}_{\cdot123}$	$r_{0\cdot123}$	$\sigma_{e_{0\cdot123}}$
$\beta_{01\cdot234}$	$\beta_{02\cdot134}$	$\beta_{03\cdot124}$	$\beta_{04\cdot123}$		$a_{00\cdot1234}$	$C_{11\cdot2}$	$C_{22\cdot1}$	$C_{33\cdot1}$	$C_{44\cdot1}$		$r_{0}^{2}{}_{\cdot1234}$	$r_{0\cdot1234}$	$\sigma_{e_{0\cdot1234}}$
$\beta_{01\cdot2345}$	$\beta_{02\cdot1345}$	$\beta_{03\cdot1245}$	$\beta_{04\cdot1235}$	$\beta_{05\cdot1234}$	$a_{00\cdot12345}$	$C_{11\cdot2}$	$C_{22\cdot1}$	$C_{33\cdot1}$	$C_{44\cdot1}$	$C_{55\cdot1}$	$r_{0}^{2}{}_{\cdot12345}$	$r_{0\cdot12345}$	$\sigma_{e_{0\cdot12345}}$

CONCLUSION

There are many additional points which should be discussed but this paper is already sufficiently long. In summing up this portion of the material, I submit the thesis that the entries of the abbreviated Doolittle solution are so useful in finding multiple and partial coefficients of correlation and regression that the standard computational formulas should be expressed in terms of these entries rather than in terms of determinants or other functions of the correlations.

REFERENCES

The bibliography given is that presented at the 103rd Annual Meeting of the American Statistical Association in joint session with the Institute of Mathematical Statistics, New York, December 28, 1941 (with slight corrections and additions) and covers a somewhat wider field than does the present paper.

1932

1. P. Horst. "A Short Method for Solving for a Coefficient of Multiple Correlation." *Annals of Math. Stat.* 3 (1932) pp. 40–44.
2. P. Horst. "A General Method of Evaluating Multiple Regression Constants." *Jour. of Am. Stat. Assn.* 27 (1932) pp. 270–278.
3. A. C. Aitken. "On the Evaluation of Determinants, the Formation of their Adjugates, . . ." *Proc. of Edinburgh Math. Soc.* Second Series, Vol. 3 (1932) pp. 207–219.

1934

4. R. A. Fisher. *Statistical Methods for Research Workers.* 5th Edition, 1934. Oliver and Boyd.
5. R. Frisch. *Statistical Confluence Analysis by Means of Complete Regression Systems.* Publication no. 5 from University Institute of Economics, Oslo, 1934.

1935

6. M. Roff. "Some Properties of the Communality in Multiple Factor Theory." *Psychometrika.* 1 (1936) pp. 1–6.
7. G. Strecker. "On Evaluating a Coefficient of Partial Correlation." *Annals of Math. Stat.* 6 (1935) pp. 143–145.
8. F. V. Waugh. "A Simplified Method of Determining Multiple Regression Constants." *Jour. of Am. Stat. Assn.* 30 (1935) pp. 694–700.

1936

9. A. K. Kurtz. "The Use of the Doolittle Method in Obtaining Related Multiple Correlation Coefficients." *Psychometrika.* 1 (1936) pp. 45–51.
10. H. D. Griffin. "A Further Simplification of the Multiple and Partial Correlation Process." *Psychometrika.* 1 (1936) pp. 219–228.
11. F. V. Waugh. "The Analysis of Regression in Subsets of Variables." *Jour. of Am. Stat. Assn.* 31 (1936) pp. 729–730.

1937

12. A. C. Aitken. "Studies in Practical Mathematics, I. The Evaluation with Applications of a Certain Triple Product Matrix." *Proc. Roy. Soc. Edinburgh.* 57 (1937) pp. 172–181.
13. P. S. Dwyer. "The Simultaneous Computation of Groups of Regression Equations and Associated Multiple Correlation Coefficients." *Annals of Math. Stat.* 8 (1937) pp. 224–231.

1938

14. W. E. Deming. "Some Notes on Least Squares." The Graduate School, Department of Agriculture, Washington, 1938.
15. L. R. Tucker. "A Method for Finding the Inverse of a Matrix." *Psychometrika.* 3 (1938) pp. 189–197.
16. F. L. Wren. "The Calculation of Partial and Multiple Coefficients of Regression and Correlation." *Jour. Ed. Psy.* 29 (1938) pp. 695–700.

1939

17. W. Ledermann. "On a Shortened Method of Estimation of Mental Factors by Regression." *Psychometrika.* 4 (1939) pp. 109–116.
18. P. S. Dwyer. "The Contribution of an Orthogonal Multiple Factor Solution to Multiple Correlation." *Psychometrika.* 4 (1939) pp. 163–171.
19. J. W. Dunlap. *Workbook in Statistical Method.* Prentice-Hall, New York, 1939.

1940

20. L. Guttman. "Multiple Rectilinear Prediction and the Resolution into Components." *Psychometrika.* 5 (1940) pp. 75–99.
21. O. Reiersøl. "A Method for Recurrent Computation of All the Principal Minors of a Determinant, and its Application to Confluence Analysis." *Annals of Math. Stat.* 11 (1940) pp. 193–198.
22. W. D. Baten. "A Solution of Normal Equations Giving the Standard Errors of the Constants." *Jour. of Ag. Research.* 61 (1940) pp. 237–240.
23. P. S. Dwyer. "The Evaluation of Multiple and Partial Correlation Coefficients from the Factorial Matrix." *Psychometrika.* 5 (1940) pp. 211–232.
24. M. Roff. "Linear Dependence in Multiple Correlation Work." *Psychometrika.* 5 (1940) pp. 295–298.
25. J. C. Flanagan. "A Successive Approximation Solution for Prediction Problems Involving a Large Number of Variables." *Proceedings of Educational Research Forum.* pp. 75–79. Published 1940 by International Business Machines Corporation.
26. W. H. Stead and others. *Occupational Counseling Technique.* American Book Co., 1940. Especially the Wherry-Doolittle Method, pp. 245–250.

1941

27. H. C. Carver. *Anthropometric Data.* Edwards Brothers, Ann Arbor, 1941.
28. P. S. Dwyer. "The Solution of Simultaneous Equations." *Psychometrika.* 6 (1941) pp. 101–129.
29. P. S. Dwyer. "The Evaluation of Determinants." *Psychometrika.* 6 (1941) pp. 191–204.

30. P. D. Crout. "A Short Method for Evaluating Determinants and Solving Systems of Linear Equations with Real or Complex Coefficients." American Institute of Electrical Engineers, 33 West 39 Street, New York City, or *Marchant Methods, MM-182*. Sept., 1941. Marchant Calculating Machine Co., Oakland, California.
31. P. G. Hoel. "On Methods of Solving Normal Equations." *Annals of Math. Stat.* 12 (1941) pp. 354–359.
32. M. Ezekiel. *Methods of Correlation Analysis*. Wiley (1941). Second Edition.
33. P. Horst (Editor). *The Prediction of Personal Adjustment*. Bulletin 48 (1941) of the Social Science Research Council. Especially the material by L. Guttman on Mathematical and Tabulation Techniques, pp. 251–364.
34. M. D. Bingham. "A New Method for Obtaining the Inverse Matrix." *Jour. of Am. Stat. Assn.* 36 (1941) pp. 530–534.
35. P. S. Dwyer. "The Doolittle Technique." *Annals of Math. Stat.* 12 (1941). pp. 449–458.
36. P. S. Dwyer. "The Evaluation of Linear Forms." *Psychometrika*. 6 (1941). pp. 355–365.
37. J. H. Smith. *Added Column Method for Computing Multiple Correlation Constants*. First Edition December 1941. University of Chicago Press.

USE OF THE DISCRIMINANT FUNCTION FOR MORE THAN TWO GROUPS

By Besse B. Day and Marion M. Sandomire
U. S. Department of Agriculture

DIFFERENTIATING between two or more populations by means of a single measurement has now become a standardized process, with well-developed fundamental theory. Conclusions thus reached, however, are likely to be inadequate. Usually, two or more measurements yield much more information than does one about the populations being studied. When two or more measurements are available, therefore, the question arises, How may these be compounded to obtain the best possible function for discriminating the populations under study? The discriminant function proposed by Fisher (4) in 1936 and further outlined in the 7th edition of his *Statistical Methods* (6) affords a procedure for obtaining the best linear function for this purpose.

This technique has been applied in widely differing fields as craniometry (1), taxonomy (4), plant selection (12), soil science (2), material psychology (13), eye color (10), and finance (3). As Fisher pointed out (7), considerable work in theory remains to be done. An analysis-of-variance test for significant discrimination of two groups of samples has been solved (5) by an adaptation of Hotelling's T^2, thus determining the comparative effectiveness of two discriminant functions. Such an exact test for a case of three or more groups with small samples is not as yet available. However, Hsu (9) has developed such a test for large samples. In addition, methods for calculating the standard errors and the significance of the individual coefficients remain undetermined.

Measurement data on Whitetail deer were generously placed at our disposal, in advance of their publication, by Barry C. Park, wildlife specialist of the U. S. Forest Service. Over a period of years, data on deer killed by hunters on the Alleghany National Forest, in Pennsylvania, had been collected under Park's direction. The data were of two types: Age classes as determined by a dental formula, and direct measurements of the body and (in the case of male deer) the antlers. As stated by Park, the dental-formula method of estimating age class, while considered accurate, is impractical for general use because only a highly trained person can apply it, and prying the jaws apart after freezing is difficult. Hence there was need for a simpler method.

The discriminant function method would seem to be more appropriate than the regression method since age (the dependent variable) can be obtained in broad classes, only, by use of the dental formula. The

present paper is concerned with the application of the discriminant function to estimating age class of deer. We shall present some intuitive methods, without proof, for overcoming the statistical limitations mentioned.

DESCRIPTION OF DATA

In the original data for age of deer taken by the dental-formula method, in the period 1936–39, four age classes were distinguished: Fawns, about 9 months old; young, about 1½ years old; mature, 2½ to 4½ years old; and overmature, 5½ years old or older. The herd studied included few very old animals. The data taken on deer killed in 1936 and some in 1937 did not include body measurements other than weight, and therefore were not used in the study except as a check on those functions obtained from weight and antler measurements. The sizes of the samples used in calculating the discriminant functions are given in Table I, the four age classes being designated by roman numerals I to IV. In the same table are listed the several characters measured, seven for males and three for females, and their mean values by sex and age class, and the standard deviations within classes. There were no significant differences in the measurements from year to year.

TABLE I

NUMBER OF DEER AND MEANS OF DEER MEASUREMENTS, BY SEX AND AGE CLASS*

Sex and measurement	9 months I	1½ years II	2½–4½ years III	5½ years and over IV	Standard deviation within age classes
Male					
Number	*105*	*64*	*179*	*154*	—
x_1, Dressed weight (lb.)	54.3	95.4	118.5	135.0	13.4
x_2, Length of body (cm.)	140.9	167.8	179.9	187.3	9.5
x_3, Length of hind foot (cm.)	42.3	47.5	49.6	50.3	1.8
x_4, Spread of antler (cm.)	—	18.0	34.2	41.5	7.1
x_5, Circumference of main beam (cm.)	—	4.8	7.5	9.0	1.3
x_6, Length of antler (cm.)	—	15.9	31.8	39.6	6.8
x_7, Number of points	—	2.9	6.5	7.7	1.7
Female					
Number	*111*	*54*	*130*	*130*	—
x_1, Dressed weight (lb.)	51.2	80.0	90.6	98.8	9.2
x_2, Length of body (cm.)	138.3	157.1	161.3	168.6	8.9
x_3, Length of hind foot (cm.)	41.5	45.7	46.1	47.5	1.9

* Supplied by Park from the original study (11).

STATISTICAL PROCEDURE

Effectiveness of single character. Despite the fact that several characters vary materially with age, considerable error might result in estimating the age of an individual deer from any one of them. Only if the measurements were average would they afford in themselves a dependa-

ble index of age. Frequency distributions showed the considerable variation in measurements for the deer in any age group (11).

If the arrays of the several measurements by age class are homoscedastic, the procedure is straightforward for the limits of the age classes fall midway between the group means. Then division of half the difference of the means for two consecutive age classes by the standard deviation within groups yields what might be considered a normal deviate. Its corresponding probability, P, found in certain tables, is the probability of an observation falling outside the range $-x$ to $+x$, where x is the deviate. In the consideration of any two successive age groups, only one tail of the probability curve is involved. Then we take one-half of P for the chance of a deer's falling either above or below the midpoint between the averages for two successive age groups, that is, above the smaller or below the larger age class. For age classes II and III, the total probability of making a mistake is the sum of two probabilities. For example, the chance of making a mistake in estimating the age of a $1\frac{1}{2}$-year deer is the probability of classing it as less than $1\frac{1}{2}$ years plus that of classing it as over $1\frac{1}{2}$ years.

The calculated probabilities for each character measured appear in Table IV under "Single measurement." For both males and females, dressed weight proved to be a better basis for estimate than either of the two other body measurements, although still not very dependable. For males, if dressed weight alone were used, there would be 6 chances in 100 of putting a 9-month deer in the $1\frac{1}{2}$-year class or above, 27 chances of misclassifying a $1\frac{1}{2}$-year, 47 for a $2\frac{1}{2}$–$4\frac{1}{2}$-year, and 26 for an overmature animal. The three antler measurements other than number of points were found to be better indicators of age than any of the three body measurements.

Compounds of two or more characters. We wish then to find the expression which will best characterize the age of individual deer, that is, the compound of measurements which will show the maximum change with age group relative to the variance of the compound within the age groups. If a linear function of the several deer measurements which shows the most distinct variation with age be represented by the linear expression

$$\phi = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_p x_p$$

then the coefficients $\lambda_1, \lambda_2, \cdots, \lambda_p$ are the solutions of the equations

$$\begin{cases} \lambda_1[x_1^2] + \lambda_2[x_1x_2] + \cdots + \lambda_p[x_1x_p] = B_1 \\ \lambda_1[x_1x_2] + \lambda_2[x_2^2] + \cdots + \lambda_p[x_2x_p] = B_2 \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad\qquad \vdots \qquad\quad \vdots \\ \lambda_1[x_1x_p] + \lambda_2[x_2x_p] + \cdots + \lambda_p[x_p^2] = B_p \end{cases}$$

where $[x_q x_p]$ is a simplified scheme for writing the pooled sums of squares and of products within the age classes. The λ's may be thought of as the weights of the variables in the function. The sums of squares and cross products (5; 6, art. 49.2) without regard to group means might have been used instead, and the resulting coefficients would have been proportional to the λ's obtained here.

The B's, which constitute the right-hand side of the equations, are the regressions of the means of the measurements on age. Thus,

$$B_p = \frac{\sum_{A=I}^{IV} (\bar{x}_{pA} - \bar{x}_p)(a_A - \bar{a})}{\sum_{A=I}^{IV} (a_A - \bar{a})^2}$$

where $\bar{x}_{pA}$ is the mean of measurement p for age A. This is a generalization of the procedure for discriminating two groups by use of the difference between the two means. The average difference among groups with respect to any one measurement is thus represented by the slope of the function of age (4, p. 185).

When the slope is to be used as a measure of the average change, it is desirable that the function of age be linear. The mean values for each character (Table I) were plotted first with rectangular coordinates, consecutive age groups being equally spaced. To overcome the nonlinear character of the relationships, a logarithmic transformation of the age values 1, 2, 3, and 4 was introduced.

Following Barnard (1) in the preliminary selection of those characters which would yield the best discriminating function, "Student's" t was calculated for a comparison of measurement values for the extreme ages, II and IV, since II was the youngest age for which all measurements were available. The resulting values for male deer appear in the first line of Table II. Next, the effect of each measurement, in turn, was eliminated from each of the other measurements, and the t values were recalculated for these adjusted measurement values.

Since the association between characters is considerable, the selection was to be based on greatest independent variation. This would consist of a combination of those characters having the greatest influence on all others and those having the least association with the others. In the case of male animals, length of antler seemed to have the greatest influence on the other measurements and length of hind foot seemed to be the most independent (Table II). After the removal of the effect of these two, the values of t for each of the other adjusted measurements

were recalculated; this process was repeated after removal of the effect of dressed weight as well. The work required for calculating the t's corresponding to all combinations of two measurements removed was prohibitive at the time.

It may be noted, in the upper portion of Table II, that removal of the effect of each of the three body measurements affected the values for the other body measurements but left those for the antler measurements little changed; and that removal of the effect of each of the four antler measurements had a marked effect on the other antler values and also on the first two body-measurement values.

TABLE II

Part 1: t values for comparison of age classes II and IV for each male deer measurement with effects of indicated measurement removed

Measurement—effect removed	Dressed weight	Length of body	Length of hind foot	Spread of antler	Circumference of main beam	Length of antler	Number of points
—	18.35	12.49	9.63	22.88	20.40	24.49	22.72
Dressed weight	—	2.22	4.30	16.80	14.10	17.92	18.24
Length of body	13.60	—	6.24	19.94	17.40	21.15	20.22
Length of hind foot	16.16	10.09	—	22.56	20.38	24.12	21.95
Spread of antler	9.81	5.64	8.97	—	6.06	10.54	12.70
Circumference of main beam	10.95	6.59	9.67	11.98	—	13.80	13.22
Length of antler	7.70	2.37	8.79	5.92	2.76	—	10.15
Number of points	12.40	7.06	7.75	12.98	8.67	13.64	—

Part 2: t values for residual regressions for each male deer measurement with effects of indicated measurements removed

Length of hind foot and antler	6.48	3.75	—	7.76	5.41	—	5.84
Dressed weight and lengths of hind foot and antler	—	1.27	—	7.01	4.46	—	7.53

As a straightforward procedure for selection, an intuitive method was tried. This consisted in ranking, in Table III, the values of t in each column of Table II, giving the smallest value the rank 1. A small value of t indicates that the measurement whose effect was removed is closely related with the measurement named at the head of the column. Hence, in Table III, the smallest horizontal total of ranks will indicate the one measurement most closely related, on an average, to all others. The results were consistent with our first interpretation with respect to length of antler and length of hind foot.

To verify our own interpretation of these t values, instead of pro-

ceeding with the development of one function only, we tested several different combinations of measurements, some of which were not promising. Since the 9-month animals were without antlers, two separate

CHART I

AGE DISCRIMINANT FUNCTION FOR MALE DEER USING THE THREE MEASUREMENTS: DRESSED WEIGHT, LENGTH OF HIND FOOT, LENGTH OF ANTLER

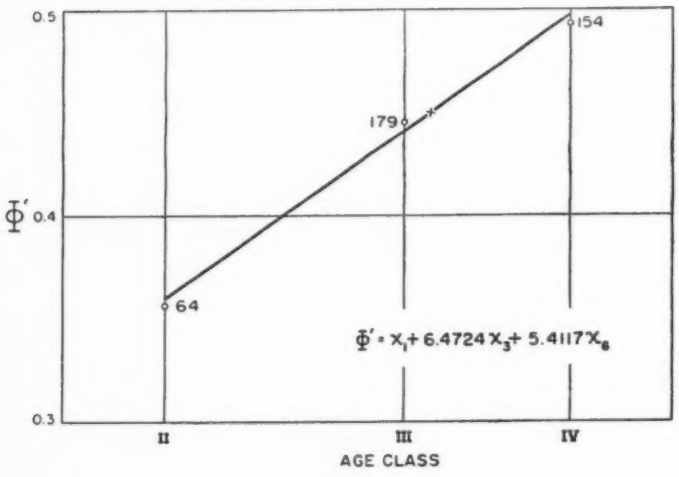


TABLE III

RANK ACCORDING TO EFFECT REMOVED OF THE t VALUES IN TABLE II

Measurement—effect removed	Dressed weight	Length of body	Length of hind foot	Spread of antler	Circumference of main beam	Length of antler	Number of points	Total of ranks for effect removed
Dressed weight	—	1	1	4	4	4	4	18
Length of body	5	—	2	5	5	5	5	27
Length of hind foot	6	6	—	6	6	6	6	36
Spread of antler	2	3	5	—	2	1	2	15
Circumference of main beam	3	4	6	2	—	3	3	21
Length of antler	1	2	4	1	1	—	1	10
Number of points	4	5	3	3	3	2	—	20

analyses were made for the male deer. One, comparable to that for the female deer, included all age groups but only the body measurements; in the other, the 9-month animals were omitted.

To determine the relative effectiveness of these functions in estimating deer age, the procedure used for one measurement was followed in

TABLE IV

PROBABILITIES OF MISCLASSIFYING DEER FOR VARIOUS MEASUREMENTS SINGLY AND IN COMBINATION AS A BASIS FOR ESTIMATING AGE, AND COMPOSITE TEST

Males of age groups, II, III, and IV[1]											All animals[2]						
Dressed weight	Length of body	Length of hind foot	Spread of antler	Circumference of main beam	Length of antler	Number of points	Probabilities, by age groups			U	Dressed weight	Length of body	Length of hind foot	Probabilities, by age groups			U
							II	III	IV					I	II	III	IV	
Male																		
Single measurement																		
*							.22	.48	.26	9.46	*			.06	.27	.47	.26	9.97
	*						.27	.62	.35	6.70		*		.08	.34	.61	.35	8.15
		*					.28	.70	.42	5.27			*	.07	.35	.70	.42	7.68
			*				.18	.45	.27	11.20								
				*			.20	.47	.27	10.96								
					*		.17	.43	.26	11.77								
						*	.20	.51	.31	9.45								
Two measurements																		
*	*						.19	.46	.27	9.89	*	*		.04	.22	.44	.26	10.97
*		*					.18	.46	.28	9.96	*		*	.03	.20	.45	.28	11.34
*			*				.10	.35	.25	13.01		*	*	.04	.27	.58	.35	9.81
*				*			.13	.38	.25	12.28								
*					*		.11	.35	.24	13.00								
	*	*					.23	.56	.33	7.60								
	*			*			.13	.40	.27	11.63								
		*	*				.09	.37	.28	12.87								
		*		*			.12	.39	.27	11.99	*	*	*	.03	.20	.45	.28	11.61
		*			*		.09	.36	.27	12.97								
			*	*			.10	.37	.27	12.59								
			*		*		.10	.37	.27	12.80								
				*	*		.11	.38	.27	12.49								
Female																		
Three measurements																		
*	*	*					.18	.46	.28	10.09	*			.06	.34	.61	.33	8.59
*		*		*			.11	.36	.25	12.65		*		.15	.56	.75	.34	5.48
*		*			*		.09	.34	.25	13.36			*	.13	.58	.81	.36	5.12
*			*		*		.09	.33	.24	13.42								
		*	*		*		.08	.34	.28	13.53								
Four measurements																		
*	*	*		*			.11	.36	.25	12.67	*	*		.06	.35	.61	.32	8.69
*	*		*		*		.09	.33	.24	13.46	*		*	.05	.34	.61	.32	8.68
*		*	*		*		.09	.34	.25	13.80		*	*	.11	.53	.75	.33	6.48
*			*	*	*		.10	.34	.24	13.48								
Five measurements																		
*		*	*	*	*		.08	.33	.25	13.86	*	*	*	.07	.38	.63	.32	8.76

[1] Total number, 397. Class I animals omitted because they have no antlers.

[2] Males, 502; females, 425.

obtaining the mean class values, normal deviates, and probabilities of misclassification. The probabilities for the various discriminant functions considered have been assembled in Table IV. These may be used for an appraisal of the several combinations as indicators of age. The first half of the table gives the results from data for the three upper age classes of males (the deer with antlers), based on 397 animals, while the second half is from data in which all four age classes were represented, 502 male and 425 female deer.

To examine one of the functions graphically, the mean value for each age class was calculated by substituting in it the corresponding measurement means from Table I. Since it is necessary to maintain only the proportionality of the λ coefficients, the calculations may be simplified by assigning to one coefficient the value unity and adjusting the others by division. One of the functions and the means are shown in Chart I plotted against the age-group values, which are maintained on a logarithmic scale. The linear representation may then be passed through the weighted means with a slope equal to the derivative of the function. Since B_p was the regression of $\bar{x}_p$ on age, the derivative is for all practical purposes equal to $\lambda_1 B_1 + \lambda_2 B_2 + \cdots$.

COMPOSITE TEST FOR RANKING DISCRIMINANT FUNCTIONS

To judge two or more discriminant functions for comparative effectiveness, it was desirable to use a single value rather than a series of probabilities. The best discrimination between two groups would be accomplished with the function that produced the largest difference between the calculated group means. When more than two groups are under consideration, an average of all the differences between successive groups is given by the slope of the function which represents the change in the function per unit change in the independent variable. Thus, for the purpose of discrimination, the maximum slope is to be desired. This, of course, was the original intent in maximizing the regression sum of squares in calculating the discriminant function. We might say, then, that of two discriminant functions the one with the greater slope is the more effective. However, the variability within classes must be taken into account if a classification of individuals is to follow. Then the change in standard units per unit change in age might be considered a measure of the discriminating effectiveness of any particular function. This is the ratio of the slope to the standard error of the function; and the larger this ratio is, the better does the compound measurement discriminate among age groups.

If two functions have equal slopes, the one with the smaller error

should be considered the more effective. This is consistent with the notion just expressed that the larger the ratio, the better is the discrimination. Therefore, we propose the intuitive test, by definition,

$$U = \frac{\text{slope of the function}}{s_\phi}$$

where the slope of the function is the derivative of ϕ and is $\sum \lambda_p B_p$. The U values are listed in Table IV for all combinations shown.

In the three-age analysis no single measurement discriminates age with any acceptable degree of precision, the failure being particularly striking for animals of age class III. Of the body measurements, dressed weight is considerably better than either of the other two. Adding one or both of the length measurements does not improve the precision of estimate perceptibly. A function of these two without dressed weight is the poorest for any combination of two.

Any one of the antler measurements, except number of points, is better than dressed weight. Length of antler is the best single measure; number of points has small value. Compounds of two antler measurements give better estimates than one alone, spread and length of antler being the best. A slightly better compound of two is one body with one antler character, preferably length or spread of antler. From the U value, the compound length of hind foot and length of antler would seem to be as effective as any including dressed weight.

The best combination of three is length of hind foot, spread and length of antler. When one body and two antler measurements are combined, no advantage is gained by including another antler measurement. Compare, for example, the values obtained by combining dressed weight and spread and length of antler with those obtained by combining these three with circumference of main beam.

We may now check the interpretation of the ranking of the t tests shown in Tables II and III. Length of antler has the smallest total of ranks, and has the largest U value. Length of antler and dressed weight were first and third in the ranking and their combination has the large U value of 13.00. This barely differs from the value 13.01 for spread of antler and dressed weight, which have rankings 2 and 3, respectively.

The highest U value for three measurements is obtained with length of antler, spread of antler, and length of hind foot, which ranked first, second, and last, respectively. It was a satisfaction to find that our intuitive selection of the hind foot was upheld when at least three measurements were involved.

The three measurements just mentioned together with dressed weight formed the best combination of four, as is judged by $U = 13.80$. These measurements, as we have seen, had the ranks 1, 2, 3, and 7.

In dealing with overmature deer the chance of misclassifying is only slightly less when a compound is used instead of any one single measurement other than length of body, length of hind foot, or number of points. The greatest advantage of the discriminant function over a single character is in age groups II and III of male deer. The quantity U or its square, may prove of value, since it is independent of any proportional shift in the λ's.

After this paper was prepared, we ran across a discussion of a similar, if not identical measure, for the comparison of two discriminant functions. This was done by David Durand in his book *Risk Elements in Consumer Instalment Financing* (3). He confined his treatment to the comparison of two groups only, with application to bad and good loans. Strangely enough he too had called his function U. It is equal to the $(\Sigma l_i a_i / \sqrt{\Sigma l_i l_j s_{ij}})$ where a_i is the difference between the means of the two groups for the ith measurement and the l's are the coefficients.

It is seen that our U is a generalization of Durand's in the case of several groups, for terms $\lambda_i B_i$ in our numerator, which was obtained from the slope of the function, are such generalizations of the terms $l_i a_i$ in Durand's. The denominator, which we felt was needed to tame down the U value in the case of great variability, is identical with his.

It might not prove a difficult matter to generalize the analysis-of-variance test as outlined by Fisher (5, p. 377) for testing the significance of a single discriminant function. As previously stated it has been done for large samples. However, we are not so much interested in the significance of a single function as we are in selecting the best from among a number of such similar functions. It would seem reasonable to suppose that the distribution of the generalized U would be found to be related to that of some generalized analysis-of-variance test as an extension of the distribution of Hotelling's T^2. It might be added here that the comparison of two discriminant functions undoubtedly involves the same difficulties as those connected with comparisons of multiple regressions. These have been discussed by both Fisher (7) and Hotelling (8).

APPLICATION OF DISCRIMINANT FUNCTIONS TO DATA NOT USED IN DEVELOPING THE FUNCTIONS

It remains to show how the results from this study may be applied to field data to determine age, although it cannot be too strongly empha-

sized that the functions derived and the conclusions reached for one herd of Whitetail deer are applicable only to data from similar herds. In addition to the data used in computing the functions, we had 407 sets of measurements with which we could test those functions not involving length of body and length of hind foot. The compound shown in Table V was applied to the measurements of the individual deer and age group limits were determined as for the single measurement. The theoretical probability of misclassification may now be compared with the proportion of deer actually misclassified in each age group. With these four measurements, calculated ratios for the animals of ages II, III, and IV are .16, .39, and .26, respectively—all slightly larger than the theoretical values, .10, .34, and .24. With the three measurements, dressed weight, spread and length of antler, the probabilities are .16, .38, and .33—again all larger than the theoretical values, .09, .33, and .24 (Table IV). With the two antler measurements only, similar differences were observed.

The totals in the last column of Table V show the number of animals that would be classed in the three age groups by use of the function alone. These frequencies may be compared with the true numbers for the three groups, 76, 261, and 70. With all functions, the number of mature deer is underestimated in favor of the two other groups.

TABLE V

APPLICATION OF THE CALCULATED DISCRIMINANT FUNCTION, $\phi = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_p x_p$, FOR DETERMINING THE AGE OF A GROUP OF MALE DEER NOT USED IN DEVELOPING THE FUNCTION

$\Phi' = x_1 + 2.3321x_4 + 5.2766x_5 + 2.0728x_6$

Theoretical probabilities of misclassification: .10, .34, .24

Theoretical age group based on Φ'	Age class						Total (theoretical age)
	II		III		IV		
	Number	*Per cent*	*Number*	*Per cent*	*Number*	*Per cent*	
IV 332.7 and above	0	0.0	49	18.8	52	74.2	101
III 249.9–332.6	12	15.8	160	61.3	16	22.9	188
II 249.8 and below	64	84.2	52	19.9	2	2.9	118
Total (dental age)	76	100.0	261	100.0	70	100.0	407

FURTHER INVESTIGATIONS NEEDED

In order to standardize the discriminant function method, it would be necessary to carry out three additional investigations. One should be made on a method of selecting the best of several measurements. If the *t* test is used, the ranking scheme should be investigated further. Another should be made on the validity of the test of significance of the

lambda coefficients, and a third for determining a test of significance of the difference between two discriminant functions in the case of more than two groups. Perhaps the distribution of the composite test, U, will be adequate for this if it is found to be some generalization of that of Hotelling's T^2.

LITERATURE CITED

1. M. M. Barnard. "The Secular Variations of Skull Characters in Four Series of Egyptian Skulls," *Ann. Eugenics* 6 (1935): 352–371.
2. Gertrude M. Cox and William P. Martin. "Use of a Discriminant Function for Differentiating Soils with Different Azotobacter Populations." *Iowa State Col. Jour. Sci.* 9 (1937): 323–331.
3. David Durand. *Risk Elements in Consumer Instalment Financing.* Natl. Bur. Econ. Res., Inc., Studies in Consumer Instalment Financing, No. 8 (1941).
4. R. A. Fisher. "The Use of Multiple Measurements in Taxonomic Problems." *Ann. Eugenics* 7 (1936): 179–188.
5. ———. "The Statistical Utilization of Multiple Measurements." *Ann. Eugenics* 8 (1938): 376–386.
6. ———. *Statistical Methods for Research Workers.* Ed. 7, rev. and enl. 1938.
7. ———. "The Precision of Discriminant Functions." *Ann. Eugenics* 10 (1940): 422–429.
8. Harold Hotelling. "The Selection of Variates for Use in Prediction with Some Comments on the Problem of Nuisance Parameters." *Ann. Math. Statis.* 11 (1940): 271–283.
9. P. L. Hsu. "On the Problem of Rank and the Limiting Distribution of Fisher's Test Function." *Ann. Eugenics* 11 (1941): 39–41.
10. Khint Maung. "Discriminant Analysis of Tocher's Eye-Colour Data for Scottish School Children." *Ann. Eugenics* 11 (1941): 64–76.
11. Barry C. Park and Besse B. Day. *A Simplified Method for Determining the Condition of Whitetail Deer Herds.* U. S. Dept. Agric. Tech. Bul. 840 (1942).
12. H. Fairfield Smith. "A Discriminant Function for Plant Selection." *Ann. Eugenics* 7 (1936): 240–250.
13. Noel Wallace and R. M. W. Travers. "A Psychometric Sociological Study of a Group of Specialty Salesmen." *Ann. Eugenics* 8 (1938): 266–302.

THE NEW WHEAT CONDITION FIGURES, BASED ON WEATHER FACTORS, FOR THE PRAIRIE PROVINCES*

By Fred H. Sanderson
Harvard University

In June, 1941, the Bureau of Statistics of the Dominion of Canada began to issue for the Prairie Provinces wheat condition figures based on weather factors. At the same time it discontinued the collection of the reports of crop correspondents for this crop. As far as is known, this is the first time that a government agency has attempted to base its official crop reports exclusively upon objective measurements of the meteorological conditions prevailing prior to and during the growing season.

This change is the third that has been effected during the comparatively short history of the condition reports issued by the Canadian crop reporting service. Up to 1918, the condition as reported by crop correspondents was expressed as a percentage of a "full" or "standard" crop, a definition similar to the one adopted by the United States Department of Agriculture. In 1919, the Dominion Bureau of Statistics adopted a presumably more precise method whereby the correspondents were asked to report condition as a percentage of the average yield of the previous ten years. From 1931 to 1940, the crop condition of 100 per cent was interpreted to be equal to the "long-time average yield per acre." The latter was first defined as the average yield for 1908 to 1930; with the recent shift to condition figures based on weather factors, the base was changed to the 1908–1940 average yield.

Where crop condition figures are based on the returns of crop correspondents, frequent changes in definition are unfortunate for several reasons. Condition figures as reported by correspondents are *opinions*, based on the appearance of the growing crop and supplemented by the farmer's experience of the general effect of the weather on yield. Like all opinions, they are affected by certain elements of psychological *bias*. The 100 per cent condition may be *defined* as a "condition promising a yield equal to the long-time average." But it does not usually follow that in practice, a 100 per cent condition is associated, on the average, with the average yield; a 50 per cent condition with one-half of the average yield, and so forth; or that a crop condition defined by 100 is equal to four times a crop condition defined by 25.

* The study upon which this paper is based was made while the author was a research associate at Harvard University and a part-time employee of the Agricultural Marketing Service, United States Department of Agriculture; it was the outgrowth of a weather-yield research project for improving statistical methodology in forecasting crop yields. The author is indebted to Mr. Joseph A. Becker, of the United States Bureau of Agricultural Economics, for reading and criticising the manuscript.

Generally speaking, crop correspondents in reporting condition show a triple bias: (1) In underestimating average yield prospects, (2) In underestimating the intensity of the year-to-year variations in yield, (3) In shying away from reporting more than a "fair" yield prospect at any time.

The tendency of crop reporters to underestimate average yield prospects is relevant only where the condition figures are defined as percentages of some definite "average" or "normal" yield. It will then appear that a reported condition of 100 would, on the average, actually indicate a probable yield considerably exceeding the "average" or "normal." If the final yield is expressed as a percentage of the long-time average and plotted along the y-axis of a system of coordinates, and condition along the x-axis, the line of average relationship would not be the diagonal defined by the points 0:0, 50:50, 100:100, etc., but would lie above this line. This downward bias of the condition reports would increase as the season advances.[1]

If the condition figures provided a correct indication of the *intensity* of the variations in yield, an increase in the condition by a given percentage would indicate, on the average, an equal *relative* increase in the yield. Hence, if condition is plotted along the horizontal and yield along the vertical axis, both being expressed as percentages of their respective means, and by using the same scale along both axes, the regression coefficient would be equal to unity; or, in other words, the slope of the regression line or line of average relationship would be 45 degrees.

For most condition-yield relationships, however, an increase of ten per cent in "condition" would indicate an increase in yield of *more* than 10 per cent, and the line of regression would meet the x-axis at an angle of more than 45 degrees.

But in addition, this bias is not constant over the whole range of prospective yields. Crop reporters would hesitate to report more than a

[1] In Saskatchewan, it was found that a condition of "100 per cent" reported at June 30 would correspond, on the average, to a 124 per cent yield (24 per cent above the average 1921–1937), whereas the same "condition" reported on July 31 would indicate a probable yield of 137 per cent. In Alberta, the corresponding figures are 109 and 121 per cent, respectively. Consequently, in comparing condition as reported, for example, at May 31 and June 30, allowance should be made for this tendency of condition figures to decline as the season advances. Such statements as the following, quoted from the official crop report of the Canadian government of July 9, 1940, may serve to illustrate this point:

"... slight declines in the condition of the wheat crops in Manitoba and Alberta were experienced during June while a more serious loss in condition was suffered in Saskatchewan. Expressed in percentages of the provincial long-time average yields per acre, Manitoba and Alberta wheat conditions both declined from 98 to 96 during June... In Saskatchewan, the wheat condition declined 5 points from 94 at May 31, this year, to 89 at June 30."

The report does not indicate that there is normally a decline in the condition figures from May 31 to June 30. The average "condition" as of May 31 during the period 1921–1937 was 94 for Saskatchewan and 98 for Alberta. The corresponding figures for June 30 are 87 and 93. Thus, in 1940, the condition actually *improved* from average to two above average in Saskatchewan, and from average to three above average in Alberta. The apparent decline in the condition figure has no real significance.

"slightly better than average" yield prospect at any time. In other words, reported condition would not increase proportionately with prospective yield. Condition tends to be a very sensitive indicator of low yield prospects, but becomes less sensitive as prospects improve.[2]

Elements of bias in condition figures similar to those described above have been observed in many countries where crop forecasts are based on numerical indications regarding the appearance of the growing crop as gathered from crop correspondents, and adequate methods to deal with this problem have been worked out.[3] The success of these methods depends to a great extent on the stability of the psychological factors involved, and, in particular, on the homogeneity of the conditions surrounding the collection of the reports. Any change in the procedure of collecting the reports or in definition may constitute a break in homogeneity and introduce unpredictable changes in the relationship between reported condition and yield. A long series of observations under homogeneous conditions is needed to eliminate bias from the reported condition figures. Any radical innovation will therefore impair the usefulness of a series of condition reports for forecasting purposes.

In a large number of countries, including the United States, condition figures as reported by crop correspondents were found to be valuable indicators of prospective yields; they have been used sometimes with remarkable success as a basis of forecasting the yield several months prior to the harvest. The Dominion Bureau of Statistics, however, came

[2] Take, for example, the case of Saskatchewan, June 30. When yield prospects were poor, an increase from 8 bushels below average to 6 bushels below average was accompanied, on the average, by an increase in "condition" as reported by crop correspondents from −53 to −16, or 37 points. An identical increase in yield when yield prospects were favorable (plus 5 to plus 7 bushels) was accompanied by an increase in reported condition from plus 14.5 to plus 16.8, or 2.3 points.

Where condition is plotted against yield per acre *harvested* (a practice resorted to when data for acreage sown are unavailable), it may be argued that the departure from linearity might possibly be due to abandonment of acreage subsequent to the condition report. It is, indeed, possible that the condition reports made by crop correspondents—particularly the earlier reports—refer more closely to acreage sown than to acreage harvested. Widespread abandonment of acreage subsequent to the report will, therefore, introduce an apparent downward bias in the condition figures when the latter are compared with final yield per *harvested* acre. As abandonment of acreage tends to be associated with low yields rather than with high yields, this apparent downward bias will be particularly noticeable in years of poor yields, while in years of above-average yields and correspondingly small extent of abandonment, it would scarcely affect condition figures at all.

If this view was more than a partial explanation of the curvilinear bias present in condition figures, the relationship between reported condition and yield per acre *sown* would be a straight line. It was found, however, that where yield data were available per seeded acre as well as per harvested acre, the curvature was somewhat attenuated in the former case, but did not disappear. A further confirmation of this is provided by the case at hand, for the Canadian yield data refer to *acreage sown*.

However, straight-line relationships between "condition" and yield may be found, especially early and late in the season. In the early part of the growing season, the total correlation between "condition" and yield is usually poor, and the curvilinearity of the relationship is frequently blurred by errors. As the season advances, the departure from linearity would become more apparent. Shortly before harvest, however, prospects of bumper crops begin to become recognized, and the relationship between "condition" and yield tends to approach a straight line.

[3] See United States Department of Agriculture, Miscellaneous Publication No. 171, *The Crop and Livestock Reporting Service of the United States* Washington, D. C., November 1933, pp. 22–32.

to the conclusion that the "condition figures obtained during the growing season are not sufficiently precise to serve as a basis for quantitative forecasting."[4]

No forecasts of production have been issued by the Canadian government since 1927. Up to 1918, official forecasts of yield and production obtained by the "par" system adopted by the United States Department of Agriculture in 1912 had been issued. The condition as reported by crop correspondents was expressed as a percentage of a "full" or "standard" crop. Forecasts of production were issued early in June, July, and August.

In 1919, at the suggestion of the International Institute of Agriculture, the Dominion Bureau of Statistics shifted to a presumably more exact method whereby the correspondents were asked to report condition as a percentage of a ten-year average. In making a forecast of yield, this reported percentage was applied directly to the average yield of the previous ten years. The reasons given for this change are as follows:

"The disadvantage of (the old) system is that the 'standard' or 'full' crop is a matter of opinion and dependent entirely upon the mental conception of each correspondent as to what such a standard should be. Consequently, the system has no true statistical basis."[5]

Within a year, this method proved unsatisfactory, and it was discontinued, though without mention that the true relation between condition figures as reported and final yields might differ from the one presupposed by the definition officially adopted.[6]

Between 1920 and 1923, the Dominion Bureau of Statistics did not publish forecasts of yield and production as of June 30. Forecasts on the basis of June 30 condition were resumed in 1924, but definitely terminated in 1927. Forecasts using condition at July 31 were continued until 1927. Since 1928, no official interpretation of condition figures has been made by the Canadian government. A preliminary estimate of production, issued early in September, is based on crop correspondents' estimates of probable yield.

Do the condition figures for spring wheat in the Prairie Provinces have value for forecasting yield per acre? To investigate this question, the condition figures as reported by correspondents during the period 1921–1937 were plotted against the final yields, separately for the two

[4] C. F. Wilson, Dominion Bureau of Statistics, Ottawa: *Relations Between Weather Factors and Wheat Yields in Western Canada.* Paper presented at the joint meeting of the Canadian Society of Technical Agriculturalists and the American Statistical Association, Ottawa, June 29, 1938. Mimeographed.

[5] Dominion of Canada, Department of Trade and Commerce, Census and Statistical Office: *Monthly Bulletin of Agricultural Statistics*, Vol. 11, No. 116, April 1918, p. 103.

[6] Canada, Dominion Bureau of Statistics: *Monthly Bulletin of Agricultural Statistics*, Vol. 13, No. 143, July 1920, p. 153.

provinces of Saskatchewan and Alberta. The significant correlation coefficients or indexes[7] of Table I were found.

TABLE I

Date of report	*Saskatchewan*	*Alberta*
May 31	0.284	not significant
June 30	0.570*	0.525*
July 31	0.839*	0.614*

* Indicates curvilinear relationship.

In view of the comparatively high correlations[8] obtained, it would seem that the Canadian condition figures as reported by crop correspondents can be used successfully as a basis of forecasting wheat production.

When the Dominion Bureau of Statistics made its recent shift from condition figures based on crop reporters' returns to condition figures based on weather factors, it announced the change with the following statement:

[7] The square of the correlation coefficient (or index) is used rather than the index itself. The squared correlation coefficient or index measures the percentage of the total year-to-year variability in yields that is associated with the predictor (or predictors).

All correlation coefficients and indexes have been corrected for bias, according to the formula

$$\overline{R}^2 = 1 - \left[(1 - R^2) \frac{n-1}{n-m} \right]$$

where n is the total number of observations and m the number of constants used. See M. Ezekiel: *Methods of Correlation Analysis*, 1930, p. 177.

These correlations were computed in 1938. Forecasts based on condition reports for 1938 and 1939 were well within the range of the errors of estimate during the period covered by the analysis. Of ten forecasts, six did not exceed the standard error of estimate; none exceeded twice the standard error. In 1940, however, the condition figures, with two exceptions, fell rather far off the mark. Throughout the Prairie Provinces, the 1940 condition figures were unusually poor indicators of the final yield. In some cases, the deviations from the line of average relationship were among the largest hitherto observed.

However, none of these deviations was larger than 1.7σ. In other words, even the largest of these errors did not exceed what one would expect to find in random sampling, at least once in about 10 cases. Moreover, in making this statement, we assume that our statistical regression curve is identical with the "true" regression. If the standard error of the regression is taken into account, we obtain, for the individual forecast, a standard error which is appreciably larger than the standard error of estimate (compare M. Ezekiel: *Methods of Correlation Analysis*, 2nd edition, 1941, p. 344).

[8] As it has been pointed out above, the definition of the 100 per cent condition was changed in 1931 from "the condition promising the average yield per acre of the previous 10 years" to "a condition promising the average yield 1908–1930."

Although strictly speaking there is no safe basis for any guess regarding the effect of this change in definition, there is a considerable question whether the slight modification may not have been overlooked entirely by most crop correspondents. In computing the above correlations, therefore, there seemed to be justification for neglecting it altogether. In fact, experience in various countries has shown that whatever the basic definition adopted by the crop reporting service, the base or 100 per cent condition in the minds of crop correspondents turns out to be a fairly stable concept, corresponding approximately, not to the long-time average, but to a long-time "normal" condition which would indicate above-average yield prospects. Crop correspondents seem to have a definite idea as to what their fields "ought" to produce in the absence of damage. In this sense, a "normal" crop is taken to be a full crop, and years of widespread or complete crop failure appear to be largely forgotten when the current crop is appraised in terms of the "normal" condition. The Canadian condition figures are no exception to this rule.

> In Table 3 a comparison is afforded between the wheat condition figures in the Prairie Provinces based on the weather factors, and those based on the returns of crop correspondents as they both relate to the condition figures corresponding with the final yields per acre for each year's crop. The three sets of condition figures employed in the comparison have been adjusted to percentages of the new 1908–1940 long-time average yields per acre for each province.
>
> A comparison of the condition figures shown indicates that on an average of 7 out of every 10 instances, the condition figures based on weather factors approximate more closely the final yield of the crop than have the previously published condition figures. In the majority of the remaining instances, where the previously employed condition figures have been more sensitive to changes in crop prospects, the reasons have been the occurrence of heavy rust or insect damage, which in turn have not been related to changes in the weather factors. In actual practice in the future, the wheat condition figures based upon the weather factors can be adjusted in the event of abnormal grasshopper activity, or in the event of rust damage, which will be less probable in the future than in the past. Accordingly, use of the condition figures based on weather factors is expected year in and year out to provide a more accurate indication of the numerical change in Prairie wheat prospects than has been available in the past.[9]

This statement fails to do justice to the usefulness of the condition figures based on the returns of crop correspondents and it is apt to give rise to false hopes as far as the new system is concerned. In fact, the above comparison seems misleading on a number of counts which cumulatively contribute to distort the picture in favor of the new system.

(1) In attempting to improve the condition figures as reported by crop correspondents, the latter were apparently adjusted to the 1908–1940 base in the following manner:

1921–1930:

$$\text{Adjusted cond.} = \text{published cond.} \times \frac{\text{Av. yield of preceding 10 years}}{\text{Av. yield 1908–1940}}.$$

1931–1940:

$$\text{Adjusted cond.} = \text{published cond.} \times \frac{\text{Av. yield 1908–1930}}{\text{Av. yield 1908–1940}}.$$

It was assumed that a change in the base would, on the average, lead to a proportionate change in the condition figures. Owing to the rather complex relationship existing between "condition" and yield, it turned out that this "adjustment" of the condition figures, instead of improving the correlation with final yield, actually resulted in lower coefficients in all cases. In Table II the correlation coefficients or indexes obtained with the unadjusted condition figures are compared with the ones obtained after adjustment.

[9] Canada, Dominion Bureau of Statistics, Agricultural Branch: *Quarterly Bulletin of Agricultural Statistics*, Vol. 34, No. 390, April–June 1941, pp. 116–119.

TABLE II

Date of report	Saskatchewan		Alberta	
	Before adj.	After adj.	Before adj.	After adj.
May 31	0.284	0.000	0.000	0.000
June 30	0.570*	0.320*	0.525*	0.505
July 31	0.839*	0.771*	0.614*	0.557*

* Indicates curvilinear relationship. Regression curves were fitted wherever this led to a statistically "significant" increase in the correlation.

The corresponding scatter diagrams appear to corroborate the hypothesis that the slight change in definition in 1931 might well have passed practically unnoticed by the crop correspondents.

(2) In the comparison made by the Dominion Bureau of Statistics of the two sets of condition figures, account was not taken of the fact, indicated above, that wherever condition figures are defined as percentages of some definite "average" or "normal" condition, they tend to fall short of actual wheat prospects.

(3) Allowance was not made for the tendency of crop correspondents —also described above—to underestimate the intensity of variations in yield.

(4) Allowance was not made for curvilinear bias present in most condition figures based on crop correspondents' returns.

Ordinary correlation methods such as are used in the United States make allowance for these systematic biases.

In the series of "condition figures" calculated from weather factors with which comparison is made, the types of bias summarized under headings (2) to (4) are not present because: (a) The method of multiple correlation starts out by equalizing the average computed yield with the average actual yield.[10] (b) The least-squares principle automatically allows for differences in the intensity of the variations in the predictor series as compared with those present in the dependent variable. The formula is so constructed that if computed yields are compared with actual yields, the positive and the negative errors just balance each other; the sum of the errors of estimation is equal to zero. (c) Wherever the relationship between weather factors and yield could not be represented adequately by a straight line, curvilinear correlation methods were applied by the Dominion Bureau of Statistics in working out the weather-crop formulae.

In order to obtain a fair comparison between the two methods, both sets of condition figures must be analyzed in the same way. In the comparison made by the Bureau, only the weather factors got the benefit

[10] Compare, however, footnote 12, second paragraph.

of an adequate statistical treatment. When correlation methods are applied in both cases, it appears that for Saskatchewan the three correlation coefficients (or indexes) computed on the basis of the old-style condition figures are all higher than those obtained with the new figures based on weather factors. In Alberta the correlations obtained with the new system are not materially better than those found on the basis of the old method.[11] The coefficients[12] (or indexes) obtained in the two cases are shown in Table III.

TABLE III

Date of report	Saskatchewan			Alberta		
	Cond. as published (Based on crop reporters' returns)	Cond. based on weather factors		Cond. as published (Based on crop reporters' returns)	Cond. based on weather factors	
	1921–1937	1921–1937	1921–1940	1921–1937	1921–1937	1921–1940
May 31	0.284	0.254	0.141	0.000	0.452	0.353
June 30	0.570	0.499	0.559	0.525	0.723	0.688
July 31	0.839	0.800	0.789	0.614	0.725	0.755

(5) A recent study[13] by the writer indicates that it is possible to improve the wheat condition figures as reported by crop correspondents for the provinces of Saskatchewan and Alberta, with the aid of one or two weather factors which crop correspondents apparently do not sufficiently consider when judging the condition of the crop. It was found

[11] Saskatchewan, the most important of the wheat-producing provinces, accounts for more than one-half of the Canadian wheat acreage; Alberta for one-third.

[12] The correlation indexes computed for "condition based on weather factors" differ from those shown in Dr. Wilson's study ("The Influence of Precipitation and Temperature on Wheat Yields in the Prairie Provinces, 1921–1940," Canada, Dominion Bureau of Statistics, *Quarterly Bulletin of Agricultural Statistics*, Vol. 34, No. 391, July-September, 1941, pp. 167–187. Cf. especially p. 187). The latter were computed after exclusion of years in which rust seriously reduced yields and, for Saskatchewan, after adjustment for damage by entomological factors in the years 1931–1939. In an appraisal of the adequacy of the weather-crop formulae developed by the Dominion Bureau of Statistics, however, errors due to "extraneous influences" must be taken into full account, since the problem of predicting the effect of these factors on yield is a particularly difficult one.

If, however, the regressions derived from the reduced number of observations are applied to the *whole* period, a discrepancy appears between the average estimated yield and the average actual yield for 1921–1940 (cf. *loc. cit.*, Table 2, cols. 7 and 8, and Table 8 cols. 7 and 8). Adjustment was made here for this discrepancy by computing linear correlations between estimated and actual yields. The correlation coefficients thus obtained were then corrected to allow for degrees of freedom. The latter were derived by subtracting the number of constants indicated in the study (compare also Table IV) from the total number of observations.

The following factors were used in Dr. Wilson's study for Saskatchewan, July 31: Seven years' weighted average precipitation, preseasonal precipitation, spring precipitation, spring temperature, summer precipitation, summer temperature. Curvilinear functions have been fitted to four of these factors.

[13] Fred H. Sanderson, Bureau of Agricultural Economics, United States Department of Agriculture: *The Use of Condition Reports and Weather Data in Forecasting the Yield per Acre of Wheat.* (Mimeographed) October 1942.

that forecasts based on official condition reports may be significantly improved by allowing for variations in subsoil moisture and for weather conditions prevailing during the period immediately preceding the report.

Lack of subsoil moisture would not necessarily affect the appearance of the crop early in the season, provided spring rains had been sufficient to germinate the crop and give it a good start. In fact, lack of subsoil moisture might not greatly affect the appearance of the crop until after heading time or later when the plant requirements for moisture are at or near their maximum, and the moisture due to precipitation during the growing season is nearly or completely used up by the plant. One might infer, therefore, that the condition of the crop in the early part of the season is influenced largely by current precipitation, and the reserve of moisture in the lower levels of the soil is neglected by crop correspondents when estimating condition. Unfortunately, no direct data regarding subsoil moisture were available for this study, and recourse had to be taken to the indirect representation of variations in soil moisture by one of its determinant factors, namely the precipitation received during the months of September, October, and November of the year preceding harvest.

It appeared likely, furthermore, that crop correspondents would fail to allow for the effect of weather conditions immediately preceding the report. As Barnes and Hopkins[14] have pointed out, even a heavy downpour does not lead to a considerable increase in soil moisture at the depth available for plant use until ten days or more after its occurrence;[15] that is, the appearance of the crop may not be affected immediately. One would expect that this lagged effect of rainfall is particularly important during the period preceding heading time (about the end of June and the beginning of July) which is considered a critical period with regard to wheat yields in the Prairie region. Consequently, it is particularly at this time of the year that any tendency of crop correspondents to neglect this factor would result in a considerable loss of reliability and timeliness in their reports.

These two hypotheses were tested statistically and it was found that the increase in the correlation coefficients or indexes resulting from the inclusion of these factors is highly significant. It would seem, in fact, that in Alberta condition taken in conjunction with the above weather factors would provide a better basis for forecasting wheat yields early in July than was possible a full month later on the basis of condition

[14] S. Barnes and E. S. Hopkins: *Soil Moisture and Crop Production*, Dominion of Canada, Department of Agriculture, Bulletin 130—New Series, Ottawa, 1930.

[15] Work with soil tensiometers at Ames (Iowa) (1938) showed that it took a week to ten days for a precipitation of two inches to penetrate 24 inches.

alone. In Saskatchewan, forecasts made early in July by means of the condition-and-weather formulae may be expected to be about as reliable as forecasts made early in August based on condition alone.

In Table IV, the corrected squared correlation coefficients (or indexes) involving condition and preseasonal precipitation are shown for

TABLE IV

Date of report	Variables used	Saskatchewan			Alberta		
		Number of constants	Squared correlation coef. or index		Number of constants	Squared correlation coef. or index	
			Uncorrected	Corrected*		Uncorrected	Corrected*
May 31	Condition as published	2	0.329	0.284	2	0.049	0.000
	"Adjusted" condition	2	0.054	0.000	2	0.034	0.000
	Condition and preseasonal precipitation	3	0.621	0.567	3	0.318	0.221
	Various weather factors (Wilson)						
	1921–1937	7	0.534	0.254	6	0.623	0.452
	1921–1940	7	0.412	0.141	6	0.523	0.353
June 30	Condition as published	3	0.624	0.570	3	0.584	0.525
	"Adjusted" condition	3	0.405	0.320	2	0.536	0.505
	Condition and preseasonal precipitation	4	0.798	0.751	4	0.702	0.633
	Condition, preseasonal precipitation, 10 days' precipitation and 10 days' temperature	6	0.907	0.865	7	0.916	0.866
	Various weather factors (Wilson)						
	1921–1937	11	0.812	0.499	8	0.844	0.723
	1921–1940	11	0.791	0.559	8	0.803	0.688
July 31	Condition as published	3	0.859	0.839	3	0.662	0.614
	"Adjusted" condition	3	0.800	0.771	3	0.612	0.557
	Condition and preseasonal precipitation	4	0.930	0.914	3	0.727	0.688
	Various weather factors (Wilson)						
	1921–1937	11	0.925	0.800	11	0.897	0.725
	1921–1940	11	0.900	0.789	11	0.884	0.755

* Corrected for number of degrees of freedom. Compare Footnote 7.

the two provinces of Saskatchewan and Alberta, for May 31, June 30, and July 31. Another index, involving, in addition, weather conditions during the ten days immediately preceding the report, is shown for June 30. It would seem that with the exception of the May 31 and July 31 reports for Alberta, the reported condition figures, when treated by correlation methods and supplemented by weather factors, may be expected to afford a better basis of forecasting wheat yields than the weather studies of the Dominion Bureau of Statistics.

(6) In comparing the new condition figures "based on weather factors" with the reported condition figures, it should be noted that the former are based on a least-squares analysis involving a considerable number of variables, whereas the latter, if treated by correlation methods, involve only two or three constants. It may be argued that the conventional correction for degrees of freedom takes care of this. There can be no doubt, however, that as the number of constants used in the analysis increases, the regression formulae derived by least-squares procedures tend to become subject to an increasing extent to the limitations and pitfalls inherent in our statistical technique. The significance of forecasting formulae tends to be reduced by such factors as the presence of joint relationships, inadequately dealt with by the conventional additive methods; the presence of serial correlation; and the appearance of spurious relationships which occur when the same body of data are used to establish a hypothesis and to prove it.[16]

This writer is well aware of the shortcomings of traditional crop reporting and shares the hope that condition figures based on correspondents' reports will be supplemented to an increasing extent or even replaced by objective methods of crop forecasting. It is in this way that progress will be made in developing the techniques for forecasting production. Dr. Wilson's analyses of the influence of weather factors on wheat yields in the Prairie Provinces represent a valuable contribution to this end. Unfortunately, the discontinuance of the collection of wheat condition reports from crop correspondents in the Prairie Provinces renders impossible any direct comparison of the relative merits of the two approaches in future years.

A REPLY

By C. F. Wilson
Dominion Bureau of Statistics

Mr. Sanderson's dominant theme is the challenging of a decision made by the Dominion Bureau of Statistics in 1941 to change the basis of its Prairie Province wheat condition figures from the one derived directly from crop correspondents' returns to a new one resulting from a study of the relations between weather and yields.[1] He asserts that the stated reasons for this decision seem "misleading on a number of counts which cumulatively contribute to distort the picture in favor of the new system." He then proceeds to elaborate six counts to which in turn this reply is addressed.

[16] In Dr. Wilson's study—at least in the case of the Manitoba and Alberta analyses—the choice and combination of the variables were guided to a considerable extent by "inspection of the scatters" (*op. cit.*, pp. 177, 181).

[1] Mr. Sanderson's footnote 12.

(1) The announcement of the change in bases carried in the June 6, 1941, crop report[2] was written for the particular group of people in the various government agencies and grain-handling organizations, in whose plans and decisions condition figures are important. By and large, by virtue of their different training, the users of condition figures are not responsive to statistical refinements. When the Dominion Bureau of Statistics decided that the weather-yield analysis provided a sufficient basis for replacing the unadjusted correspondents' condition figures with which dissatisfaction had become general, there still remained the problem of presenting the new series to its users.

The simplest procedure was to compare the correspondents' condition figures and the weather-yield results, logically those for July 31, with the final yields per acre. To facilitate direct comparisons, all three series had to be adjusted to a common base. Since the new condition figures from the weather-yield analysis were to be published in terms of the 1908–1940 base, the latter was selected. Mr. Sanderson has shown that his regressions of yield on correspondents' condition suffer some impairment when the latter are shifted to the 1908–1940 base. The writer was not, as Mr. Sanderson feels he should have been, comparing the weather-yield condition figures with the results of such regressions. The merits of the latter will be discussed under the three subsequent counts. Meanwhile, to avoid any suggestion of having affected the comparison to the disadvantage of the correspondents' condition figures by shifting them to the 1908–1940 base, direct comparison of the three series can just as readily be made in terms of 1908–1930 base.[3] In this way the condition figures *as published*[4] may be compared directly with the condition based on weather factors and with final yield, when the latter two series are adjusted to the 1908–1930 base. When such comparison is made the results if anything show an improvement over the stated 7:3 ratio in favor of the weather-yield analysis at July 31. It, therefore, cannot be implied that the users of condition figures were misled by the comparison presented between the results of the weather yield study and those of the condition figures hitherto employed.

(2), (3) and (4). These counts are like in kind in that they charge failure to take into account three types of bias present in crop correspondents' returns. Since the procedure Mr. Sanderson follows does not attempt to isolate the separate effects of the three biases, the latter

[2] Mr. Sanderson's footnote 9.

[3] The average yields per acre for the two bases are as follows:

Base	*Manitoba*	*Saskatchewan*	*Alberta*
1908–1930	16.8	16.3	17.9
1908–1940	16	15	18

[4] The published condition figures for 1931–1940 were based on the 1908–1930 average yield. While those for 1921–1930 were based on the average of the preceding 10 years, Mr. Sanderson doubts that any adjustment is warranted in comparing them with those issued on the 1908–1930 base.

may be considered here as one problem. His method is the one commonly employed by the Division of Agricultural Statistics in the United States Department of Agriculture in interpreting their condition figures, and consists of studying the regression of final yield on the monthly condition figures reported by correspondents over a period of several months preceding harvest.

The writer investigated this method of adjusting for correspondents' bias prior to embarking on the weather-yield analysis, but found the approach wanting in several respects as applied to the western Canadian situation. The reasons for its shortcomings in that particular situation were not hard to find, and were primarily concerned with the underlying hypotheses. In contrast with the winter wheat crop in the United States, the spring wheat crop in western Canada is grown within a very short season. In addition, the year-to-year yield-per-acre variability is considerably more than double that of the United States crop. While the weather determinants of yield for the United States winter wheat crop are spread over a nine-month period, the principal determinants of yield in western Canada occur within a two-month period. In the latter area, preseasonal and spring developments to May 31 play but a minor role in the final determination of yields. The crucial developments come in June and July. Even if May seeding has been done in a practically dry seed-bed, normal June rains can bring the plant into the shot-blade stage, and if July precipitation is good, a bumper yield can result. If, on the other hand, temperatures are extreme, the plant may burn or ripen prematurely; if July rainfall is light, reduced yields are the result.

Because of the brevity of the growing season, and the very wide amplitude in yields resulting from weather developments within that period, it follows that it is an invalid operation to employ the regression of yield on the reported May 31 condition as a means of isolating and adjusting for the bias in the correspondents' returns. The latter report on current condition, not on their forecasts of final yield. In the absence of any bias, one would expect at most to find but a very poor relation between final yield and "true" May 31 condition. The major determinants of yield have yet to occur. What Mr. Sanderson has isolated by his regression of yield on the reported May 31 condition is actually an inseparable admixture of correspondents' bias and the real discrepancies between final yield and "true" May 31 condition, of which the latter are of much the greater magnitude at that date.

A similar regression of yield on the June 30 reported condition is open to the same objection in principle, but modified by the fact that some of the more important determinants of yield are then on record. Even with normal June developments, however, July remains a crucial

month. Only a partial relation could be expected between yield and "true" June 30 condition, and again the regression would isolate an admixture of bias and real discrepancies.[5] At July 31, however, the major determinants are on record and it would then be a logical procedure to employ the regression of yield on condition as a means of adjusting for bias.

For the foregoing reasons the writer rejected the methods now proposed by Mr. Sanderson as a means of improving on the old-style condition figures for May 31 and for June 30. For the same reasons the writer is not inclined to attach significance to the correlation coefficients or indexes calculated by Mr. Sanderson for the correspondents' condition figures for those months. The only point remaining at issue is whether the July 31 condition figures afford, with proper adjustment, a better basis for approximating final yield than did the condition figures based on weather factors to which the change was actually made.

As evidence on this point, Mr. Sanderson makes a comparison for two of the three Prairie Provinces, on the basis of squared correlation indexes which he has calculated both for the July 31 condition based on correspondents' returns and for the July 31 condition based on weather factors. His comparison shows better results for the correspondents' returns in Saskatchewan and for the weather factors in Alberta. The squared correlation indexes he shows for both series, however, are open to some question.

In the first place, in developing the weather yield analysis, data for four years in Manitoba and one year each in Saskatchewan and Alberta were excluded because extraneous factors such as rust were dominant in those years. Their inclusion would have vitiated an approximation to the "true" net regressions of yield on the various weather factors. From that viewpoint, their exclusion was logical, but from Mr. Sanderson's viewpoint, it is necessary to compare the results of the weather analysis in *every* year of the 1921–1937 period for which he is showing his interpretation of the crop correspondents' returns.[6] The squared correlation indexes he calculated for the weather analysis implied that no adjustment whatever in practice would be made in years when extraneous factors become dominant. A more reasonable assumption would be that on the basis of the Bureau's numerous contacts with

[5] If an approximation to correspondents' bias at May 31 and June 30 were desired one might well study the differences between the correspondents' condition figures for these dates and the corresponding condition figures furnished by the weather-yield analysis. The latter are based on the partial regressions of yield on preseasonal and spring weather factors and, with allowance for their standard errors of estimate, are approximations to the "true" condition for these dates. The extent to which correlation is lacking between yield and "true" condition as of May 31 and June 30 emphasizes the logical futility of employing such data as a forecasting device. That the U.S.D.A. can use their data for forecasting purposes is undoubtedly due to the fact that their crop determinants are not nearly as heavily concentrated in the two months prior to harvest. The "true" condition figures under such circumstances would more closely approximate the final yield.

[6] See footnote 12 of Mr. Sanderson's article.

actual field developments, arbitrary adjustments would be made as close on the average as the standard error of estimate to the true adjustment for those exceptional occasions. From theoretical considerations it would follow that substitution of the standard error of estimate for the "nonsense" observations in the excluded years would leave the correlation indexes practically unchanged from those which were calculated after exclusion of the few years in which non-weather factors were dominant. On these grounds the writer rejects the July 31 correlation indexes for the weather analysis which Mr. Sanderson has calculated, and will use in comparison below the indexes as published.[7]

Secondly, just as Mr. Sanderson has asked that the years of "extraneous influence" be taken into account, he should be willing to consider the 1921–1940 period for which the data are available, rather than cutting off at 1937. The comparison shown here accordingly includes the correlation indexes calculated for the whole of the 1921–1940 period.[8]

Thirdly, one wonders why Mr. Sanderson has excluded the Manitoba case from consideration. It isn't sufficient to say that Manitoba accounts for only one-tenth of the Canadian wheat acreage. The problem of a proper basis for the publication of condition figures is just as relevant to the Manitoba situation as it is to that of the other two provinces. With Manitoba included, the comparison follows:

Period 1921–1940	Squared correlation indexes (corrected)		
	Manitoba	Saskatchewan	Alberta
July 31 condition based on correspondents' returns	0.363	0.767	0.608
July 31 condition based on weather factors	0.651	0.865	0.876

Surely there is no evidence here that the users of condition figures were misled by the adoption of the weather-yield approach, instead of adjusting for correspondents' bias in the old series in the way Mr. Sanderson has suggested.

(5) Mr. Sanderson's argument on this count is that if the Bureau had superimposed on the regression of yield on the old-style condition figures a further adjustment for certain weather data, as he has done in a forthcoming study,[9] results superior to those of the Bureau's weather analysis could have been obtained for forecasting purposes. The first objection to this procedure is the one already made against the use of regression of yield on condition at the May 31 and June 30 dates. The residuals against which the weather data would be plotted

[7] *Quarterly Bulletin of Agricultural Statistics*, King's Printer, Ottawa, Vol. 34, 1941, p. 187.

[8] In calculating the squared correlation indexes of the regressions of yield on correspondents' condition, the writer did not have access to Mr. Sanderson's regressions for the 1921–1937 period. The writer's curve for the 1921–1940 period, however, gave the same results for 1921–1937 as those shown by Mr. Sanderson.

[9] The writer greatly appreciates the effort made by the Division of Agricultural Statistics, U.S.D.A., to place in his hands a manuscript copy of Mr. Sanderson's study referred to in the latter's footnote 13

are not a valid approximation to the correspondents' bias for which the weather data are employed to correct. Secondly, one wonders at the particular choice of weather factors Mr. Sanderson has employed. Preseasonal precipitation in conjunction with May 31 condition appears logical, in conjunction with June 30 condition subject to some question, and in conjunction with July 31 condition quite illogical since correspondents at that date would have visual evidence of any significant differences preseasonal precipitation has made. Ten days' preceding precipitation and temperature, if not highly intercorrelated for such a short period is a logical choice for June 30, but it would be just as logical for July 31. That the variables as they were employed have led to "statistically" significant improvements for the period under review holds questionable promise for their use as a forecasting device when the causal relations are not thoroughly established. Thirdly, the fact that higher correlation indexes have been adduced for June 30 than for July 31 in itself points to weakness in the underlying hypotheses.

(6) The inference conveyed by Mr. Sanderson's footnote 16 is misleading and unfortunate. In the study cited, the data were employed to establish none of the major hypotheses in the analyses for any of the three provinces, such as the partial regressions of yield on June and July precipitation and temperatures, nor in the case of April and May precipitation and temperatures in the case of Saskatchewan. When the analyses for Manitoba and Alberta were undertaken, inspection of data for these early factors was made in determining whether their inclusion was warranted. To this minor extent the data were used for both hypothesis and proof. Mr. Sanderson himself appears to have made similar use of data in choosing between linear and curvilinear regressions[10] and in some of the particular choices of weather variables employed in his forthcoming study, *à propos* the selection of which reference has already been made.

Mr. Sanderson's remarks about the potential pitfalls of multiple correlation analysis when several variables are used are of general application to the method. When these considerations are addressed to the particular application of the method in the Bureau's weather-yield analysis, it should be conceded that some of the common difficulties are avoided. The problem is one of physical rather than economic phenomena, and the causal relations are clear. Among the variables used, the intercorrelations are accidental and predominantly insignificant.

EDITOR'S NOTE.—It is anticipated that this discussion will be continued in the forthcoming March issue.

[10] See Mr. Sanderson's Table II.

DISTRIBUTION OF INCOME IN 1935–36

By Rufus S. Tucker
General Motors Corporation

In two previous articles[1] the present writer has pointed out some deficiencies in the National Resources Committee's reports on the distribution of income in the United States in 1935–36. Every criticism made in those articles has been confirmed by subsequently published information, but the tentative revision suggested in the first article has turned out to be insufficiently drastic. Consequently the distribution here presented has been prepared, making use of the finally published results of the Consumers' Purchase Study, the detailed *Supplements to Statistics of Income for 1936*, and the reports of the Wisconsin and Delaware state income taxes for that year. Delaware and Wisconsin are the only states that have income tax statistics in usable form; one has a large proportion of large incomes, the other a small one.

Families. The NRC distribution for families over $7,500 was based on *Statistics of Income for 1935*, with certain assumptions as to evasion, understatement, capital gains and losses, the effect of separate reporting by married couples, and the shift in income between 1935 and 1936. The *Statistics of Income for 1936* and the special *Supplements* published for that year permit an accurate evaluation of separate reporting, and capital gains and losses, and the effect of adding back legally allowable deductions, and enable the analysis of income tax reports to be carried down to families with $3,500 income. This analysis can be readily converted to the 1935–36 period by comparison of the regular *Statistics of Income* for the two years.

If the NRC allowance for tax evasion and understatement was not excessive, and if the replacement of allowable deductions was justified, and the allowance for imputed income approximately correct, the NRC figure of 471,000 for the total number of families above $7,500 can be supported by the income tax data. But if the same method had been applied to all tax returns, instead of only to those with statutory net incomes of $5,000 or over, the number of families with economic incomes above $3,500 would have been 2,080,000 instead of the NRC figure of 1,635,000. Undoubtedly 2,080,000 is too large a figure, not being supported by any other evidence; whence it follows that 471,000 is probably too high an estimate for families above $7,500. Certainly there was no less tax evasion and probably no less understatement, in proportion

[1] *Review of Economic Statistics*, November 1940 (Vol. XXII, pp. 165–182) and February 1942 (Vol. XXIV, pp. 9–21).

to actual incomes, by taxpayers with statutory net incomes under $5,000 (or economic incomes under $7,500) than by those with larger incomes.

In each statutory net income class under $10,000 the assumption that 25 per cent of the citizens failed to file returns, and that the average returns filed understated net income by 15 per cent, though startling, might possibly be justified if taken by itself, as a measure to adjust statutory net income. But applied to total income, as it was by the NRC, it involves a double allowance for understatement. Perhaps the easiest and commonest way to understate taxable net income is to exaggerate the allowable deductions. By restoring these deductions the NRC automatically wiped out a large part of the understatement. In addition, by restoring all deductions the NRC refused to recognize that a large part of them were legitimate business expenses, which must be deducted to ascertain the correct amount of economic income. Consequently the number of incomes in the income classes above $7,500 was greatly exaggerated.

If the federal tax returns are adjusted for capital gains and tax-exempt interest and further adjusted by allowing for evasion and understatement on the NRC assumption, but without adding back statutory deductions, we get 380,000 for the number of families over $7,500 and 1,693,000 for the number over $3,500. These figures also include an allowance for imputed income, based on the average amount of imputed money value of current expenditures reported for each income class in Table 1 of *Family Expenditures in the United States.*

If no separate allowance is made for tax evasion and understatement, but total income is taken as the base instead of statutory net income, the number of families over $7,500 would be 300,000 and the number over $3,500 would be 1,139,000. These figures also include adjustments for capital gains and losses and tax-exempt income and non-money income.

The Delaware state tax returns, if adapted to the national population, would warrant taking 1,736,000 or more as the proper total for families over $3,500 without requiring any allowance for evasion or understatement, but Delaware has a greater proportion of large incomes than most states. The Wisconsin state tax returns would require a considerable allowance for evasion to come anywhere near that figure. Some such allowance must certainly be made, and it is almost as certain that the allowance should be greater for small incomes than for large ones. Consequently the figure of 380,000 has been preferred for $7,500 and over, and the figure of 1,711,000 for $3,500 and over.

This figure, 1,711,000, is based on the Consumers' Purchase Study,

TABLE I

DISTRIBUTION OF NON-RELIEF FAMILIES BY INCOME CLASSES 1935–36

Income class	Per cent As stated by NRC* (Weighted averages of localities) (233,000 cases)	Per cent As shown by CPS samples (Unweighted pool of all cases reported) (259,000 cases)
Under 250	2.83	2.13
250– 500	7.83	5.41
500– 750	11.31	8.59
750–1,000	13.40	11.33
1,000–1,250	13.16	12.77
1,250–1,500	10.80	10.98
1 500–1,750	9.05	9.98
1,750–2,000	7.34	9.15
2,000–2,250	5.46	6.97
2,250–2,500	4.04	5.20
2,500–3,000	5.23	6.66
3,000–3,500	2.99	3.96
3,500–4,000	1.76	2.21
4,000–4,500	1.00	1.55
4,500–5,000	.61	.77
5,000–7,500	1.30	1.66
7,500 and over	1.89†	.68

* Calculated from *Consumer Incomes*, p. 97, Table 9B.
† Adjusted by NRC on basis of income tax returns.

which was used by the NRC before the results of the survey were available to the general public.[2] The NRC used only part of the available material, and weighted it in accordance with various assumptions, some of which the present writer questioned in his previous articles. Table I has been prepared to show to what an extent the processing of the material by the NRC resulted in a different picture from that presented by the original data. The chief reasons for the differences were: (1) excessive weighting of the samples from the South, especially South Carolina and Georgia; (2) omission of some samples from more prosperous regions; (3) excessive weighting of broken families, with no distinction between those lacking a wife and those lacking a husband.

Both the unweighted CPS pattern and the NRC weighted pattern suffer from the non-representative nature of the communities and families sampled, although not to the same extent. Both also probably suffer from an inadequate allowance for non-money income of farmers and servants and from wide-spread understatement of income by the families interviewed, but no correction is possible on those accounts.

[2] The results of the Study of Consumer Purchases are now available in *Bureau of Labor Statistics Bulletins* Nos. 642–649 and *Department of Agriculture Miscellaneous Publications* Nos. 339, 345, 356, 370, 375, 383, 396, 399, 402, 405, 415, 422, 428, 432, 436, 452, 455, 456, 457, 462, 464, 465.

Since the published CPS samples refer only to non-relief families it is necessary to prorate them against the 24,913,000 non-relief families reported by the NRC, and to add to each class the NRC estimates of families on relief. The totals thus obtained are shown in Table II as cumulations for all classes from zero to $3,500. Above $3,500 the cumulations are based on the income-tax analysis.

TABLE II

DISTRIBUTION OF FAMILIES AND SINGLE INDIVIDUALS BY INCOME-CLASSES IN 1935–36 (cumulated)

(in thousands of units)

Economic income	Families		Single individuals		Consumer units	
	NRC	Revised	NRC	Revised	NRC	Revised
Total	29,400	29,400	10,058	10,058	39,458	39,458
Over $500	25,222	25,997	7,525	7,760	32,747	33,757
" 750	21,423	22,876	5,553	5,926	26,976	28,802
" 1,000	17,146	19,114	3,954	4,384	21,100	23,498
" 1,250	13,263	15,328	2,845	2,860	16,108	18,188
" 1,500	10,398	12,417	1,967	1,967	12,365	14,384
" 1,750	8,054	9,843	1,421	1,432	9,475	11,275
" 2 000	6,157	7,496	1,022	1,033	7,179	8,529
" 2,250	4,737	5,700	738	749	5,475	6,449
" 2,500	3,693	4,368	528	539	4,221	4,907
" 3,000	2,378	2,698	367	361	2,745	3,059
" 3,500	1,635	1,711	258	252	1,893	1,963
" 4,000	1,196	1,217	194	188	1,390	1,405
" 4,500	946	1,006	158	158	1,104	1,164
" 5,000	794	817	133	133	927	950
" 7,500	471	380	76	65	547	445
" 10,000	284	228	47	41	331	269
" 15,000	152	111	26	22	178	133
" 20,000	93	68	17	14	110	82
Median	$1,160	$1,300	$815	$875	$1,070	$1,170
Lower tercile	$840	$970	$605	$645	$780	$870
Upper tercile	$1,570	$1,755	$1,120	$1,150	$1,450	$1,580
Highest decile	$2,750	$2,900	$2,010	$2,030	$2,600	$2,725
Highest percentile	$9,800	$8,650	$6,100	$5,900	$9,100	$8,100

As a check on the figures obtained by adding the CPS samples the Delaware and Wisconsin state tax returns for 1936 can be used. The federal tax returns are of no value for families below $3,500. The Wisconsin returns are also incomplete below $1,750, since returns were not required for statutory net incomes under $1,600.

These state returns, if adjusted by the methods used by the NRC for federal returns in the higher brackets would both show a greater proportion of families above $2,250 than the unweighted samples and a greater proportion above $1,750 than the NRC distribution. The NRC

allowance for evasion and understatement would probably be excessive for families in income classes over $2,000 in those states, where the presence of both federal and state collectors and auditors might be expected to reduce evasion to less than the national average. Likewise family incomes in 1936 were doubtless somewhat higher than in 1935–36. Nevertheless with very moderate allowances for evasion and no allowances for understatement the Wisconsin returns, although incomplete below $1,750, would indicate that in 1935–36 the number of families between $1,500 and $2,250 was greater than shown by using the samples, and the Delaware returns would indicate a greater number between $1,250 and $1,750, and between $2,000 and $2,250, and larger cumulative totals at every point above $1,250.

Taken together the state tax returns show plainly that the NRC estimate of families between $1,250 and $3,000 was too low and that the unweighted samples do not increase their number unduly.

Single persons. The distribution of single persons over $3,000 as reported by the NRC was based on federal income tax returns for 1935 with adjustments for families. The same objections therefore apply. In addition the allowance for imputed income was not stated in the NRC reports, because it was "not considered sufficiently reliable to warrant presentation."[2]

The revised figures used herewith for classes over $10,000 are based on the 1936 *Supplements*, adjusted to the average of 1935 and 1936 by the ratio of the number of returns in each class as reported in *Statistics of Income* for those years. They represent the higher of two figures, one of which is based on total income adjusted for capital gains and losses and tax-exempt interest; the other being statutory net income adjusted for evasion and understatement in accordance with the NRC formula, with tax-exempt interest added. Since evasion was doubtless greater among individuals than among families the number obtained by this method between $5,000 and $10,000 has been raised to conform to the respective number of families and individuals in that class reported in the *Supplement*.

For classes between $1,500 and $5,000 the NRC figures have been retained, with a little smoothing to make them compare better to the number of families in each class. Between $1,500 and $3,000 the NRC figures should perhaps be raised. They were almost entirely derived from samples relating to women only, and the NRC's adjustment to cover men was entirely a matter of judgment. Moreover the samples were mainly collected in the years 1931–34, the bottom years of the

[2] *Consumer Expenditures*, p. 147.

depression, and no adjustment appears to have been made to bring them into line with the higher wages and greater employment of 1935–36.

Between $1,000 and $1,500, the NRC figures must certainly be raised. They are too low by comparison with either the Delaware or Wisconsin tax returns if those returns are adjusted in the same way that the NRC adjusted the federal tax returns in the higher brackets. The numbers in each class between $750 and $1,500 were too low according to both states' returns if the number reported in each class is multiplied by a factor that makes the total above $1,500 equal to 1,967,000, the NRC figure. Finally the numbers between $1,000 and $1,500 were too low to be consistent with the number of families in those income classes, as revised. Obviously if the total number of single individuals is one-third of the number of families, the proportion must be greater than one-third in the lower income classes, and the proportions between the cumulated totals must approach one-third as the point of cumulation approaches zero. The federal and state tax returns, and the NRC published estimates, all agree that the number of single individuals increases faster than the number of families as one goes down from $2,500 or $2,000 to zero incomes. Taking the ratios of families to individuals shown by the NRC estimates, which are higher than those shown by other sources and therefore give a minimum number of single individuals, the increases in the aggregates of families over $500, $750, $1,000 and $1,250 already demonstrated require that the number of individuals in the corresponding aggregates be increased as shown in Table II. This results in increasing the number in each class between $1,000 and $1,500 and decreasing the number in each class below $1,000. If the state tax returns alone had been used as basis for revision, with very moderate adjustments for evasion the class $750–$1,000 would also have been increased and the cumulative total over $750 would have been larger than shown in Table II.

Classes below $500. No attempt is here made to subdivide the families or single individuals below $500. The CPS samples in this area are open to grave suspicion. In the Consumer Expenditures volume[4] the NRC rejected flatly the statements of non-relief families and single individuals in this class concerning their expenditures. If their reported expenditures could not be relied upon, were their reported incomes reliable? Moreover it is of little importance whether the incomes of non-relief families and single individuals under $500 were above or below $250 since the overwhelming majority of non-relief families and single persons in that class in any given year are there only temporarily, and in many cases in a technical sense only, because their business losses

[4] *Consumer Expenditures*, pp. 134, 146.

and deductible expenses happen to be abnormally large. They live on their capital or on borrowed money until their income picks up, and if it does not pick up they go on relief. If, as the NRC seemed to imply in its expenditure study, a family cannot live on less than $500 nor an individual on less than $250, using the NRC definition of income, there can obviously be no families or individuals with incomes less than $500 or $250, unless they have liquid assets or good credit. The NRC definition of income included gifts and inherited cash if spent for current living expenses; also non-money income from owned homes and home-grown food and fuel, and payments in kind, and public relief in cash or in kind. And since relief is not ordinarily granted to persons with enough liquid assets or good credit to live on, and is always granted to such an extent as to prevent starvation and nakedness, it would follow that most of the 1,523,000 relief families with incomes under $500 and the 329,000 single men and 169,000 single women on relief with incomes under $250 reported by the NRC either did not exist or else should have been placed in some higher class.

From the foregoing it appears that the total of 3,403,000 families and 2,298,000 single individuals left in the under-$500 class is probably too great, but there is no substantial basis of fact to warrant any specific lower figure. It is at any rate apparent that the NRC put at least 1,000,000 and possibly 2,500,000 too many consumer units into the class with incomes under $500.

The pattern of distribution of families published by the NRC, although purporting to be based on the Study of Consumer Purchases and the income-tax statistics, differed widely from its sources and reflected more the result of various assumptions and conjectures. It greatly exaggerated the numbers over $7,500 and under $1,250. The tabulation here presented conforms more closely to the original data, but may still exaggerate the numbers over $7,500 and under $1,250 and especially under $500.

The NRC pattern of distribution of single individuals in 1935–36 was based on inadequate data, largely collected in years less prosperous than 1935–36. Checked against federal and state income tax data it appears to have overstated the numbers over $7,500 and below $750. The present tabulation perhaps still does so, but to a lesser extent.

The NRC pattern of consumer units is so far removed from the original data and from the probable facts that it is highly unlikely that any conclusions drawn from it, unless adequately supported by other evidence, can be sound. For the sake of scientific truth and sound public policy, its further use by economists should be avoided, and the many studies already published relying on it should be revised.

A PUNCHED CARD TECHNIQUE TO OBTAIN COEFFICIENTS OF ORTHOGONAL POLYNOMIALS

By Lila F. Knudsen[1]
Food and Drug Administration

In past years many methods have been developed for fitting polynomials. Although the use of orthogonal polynomials for fitting continuous functions probably originated with Legendre, Tchebycheff[2] seems to be the first to use orthogonal polynomials for fitting discrete observations with and without equal intervals. Orthogonal polynomials have been the subject of many mathematical dissertations and an entire volume[3] of references and cross references on this subject has been compiled by a committee, of which J. Shohat is chairman. The fitting of polynomials by the use of orthogonal functions has been mainly by two methods: (1) summation; and, (2) multiplication of the variates by actual values of the orthogonal polynomials as given in a table. Pearson[4] and Isserlis[5] dealt with the fitting of non-equidistant and unequally weighted data. R. A. Fisher developed methods of calculation of Tchebycheff polynomials by successive summation up to the 5th degree,[6] and in his classic work on "Yield of Wheat at Rothamsted"[7] refers to work by Esscher.[8] F. E. Allan published a paper, "The General Form of Orthogonal Polynomials for Simple Series."[9] C. Jordan[10] develops the theory in full, and crediting the method to Tchetverikoff in a paper published in Russian in 1926, demonstrates the building up of general expressions for the orthogonal polynomials in factorial form originally used by G. F. Hardy in "Graduation of British Offices Tables, 1863–93." Jordan also suggests the numerical method of building up the polynomials by summation and furnishes tables for this purpose. J. Shohat gave the mathematical approach in "Stieltje's Integrals in Mathematical Statistics."[11] A. C. Aitken[12] gives a relatively simple demonstration of the theory and includes a series of tables up

[1] The author wishes to thank Dr. Harold Hotelling of Columbia University for his helpful suggestions and criticisms, and Mr. Meyer A. Girshick of the Bureau of Agricultural Economics in the Department of Agriculture for his continued encouragement, suggestions and assistance.

[2] P. L. Tchebycheff, 1854–1875, see *Oeuvres* (1899).

[3] J. Shohat (Chairman) "Bibliography on Orthogonal Polynomials," assembled by Committee of National Research Council, *Bulletin of the National Research Council*, No. 103, August, 1940.

[4] K. Pearson, *Biometrika*, Vol. 13, pp. 296–299 (1921).

[5] L. Isserlis, *Biometrika*, Vol. 19, pp. 87–93 (1929).

[6] R. A. Fisher, *Journal of Agricultural Sciences*, Vol. 9, pp. 107–135 (1921).

[7] R. A. Fisher, *Philosophical Transactions of the Royal Society*, B 213, pp. 89–142 (1924).

[8] F. Esscher, "Uber die Sterblichkeit in Schweden (1886–1914)," "Meddelander från Lunds Astronomiska Observatorium," 2, 23, pp. 10–21 (1920).

[9] F. E. Allan, *Proceedings of the Royal Society of Edinburgh*, Vol. 50, p. 310 (1929–30).

[10] C. Jordan, *Annals of Mathematical Statistics*, Vol. 3, pp. 257–357 (1932).

[11] J. Shohat, *Annals of Mathematical Statistics*, Vol. 1, pp. 73–94 (1930).

[12] A. C. Aitken, *Proceedings of the Royal Society of Edinburgh*, Vol. 53, pp. 54–78 (1933).

to the 5th degree for the summation method. Aitken, Fisher, and F. F. Stephan[13] show individually the advantage of dividing a series of data and summing from the extremes. Two excellent resumés of the subject of orthogonal polynomials with many references are given by Wishart[14] and Irwin.[15]

Usually the objection raised to fitting orthogonal polynomials to a large mass of data is the laboriousness of the task. This objection may be overcome by using punch cards and Hollerith machines, namely, card punch, sorter, tabulator and summary punch. This article will demonstrate fitting of orthogonal polynomials to weekly weighings of rats by means of tabulating equipment, and will discuss the analysis of the coefficients of the polynomials in order to investigate the effects of animal diets.

A process of fitting orthogonal polynomials by a punched card method of successive summations has been published by Richard Warren of Columbia University.[16] However, his method involves at least as many sets of cards and as many runs as the degree of the polynomial. For instance, for fitting a polynomial of the sixth degree, six sets of cards would be used, and six runs besides the three or four runs for calculating sums of squares. This method seems cumbersome and lengthy for fitting polynomials simultaneously to more than one series of data.

The method described in this article uses the direct method of multiplying by the tabled actual values of the orthogonal polynomials and is relatively short involving only about one-fifth the time and work required by Mr. Warren's method. The author is not aware of any previous description of fitting orthogonal polynomials by the direct method of multiplication by addition on a punched card accounting machine using the digiting process.

The procedure is given here not only because of its use in fitting orthogonal polynomials but also because of its adaptability to other problems wherein values obtained from tables are multiplied by one or more variates in order to obtain some type of curve. The method has recently been used by Miss Dorothy Nickerson, Cotton Branch, Agricultural Marketing Administration, to obtain tristimulus values for 425 Munsell colors for each of four illuminants using the method described in Hardy's *Handbook*.[17] The entire process of punching, tabula-

[13] F. F. Stephan, this Journal, Vol. 27, pp. 413–423 (1932).

[14] J. Wishart, *Journal of the Royal Statistical Society*, Vol. 96, pp. 487–491 (1933).

[15] J. O. Irwin, *Journal of the Royal Statistical Society*, Vol. 97, p. 134 (1934).

[16] R. Warren, "Hollerith Machine Computation of Least Square Trend Lines," Columbia University Statistical Bureau Document #2 (1933).

[17] A. C. Hardy, *Handbook of Colorimetry*, Massachusetts Institute of Technology (1936).

tion and calculation of the tristimulus values for each of the four illuminants from the 425 spectrophotometric curves (one for each of the Munsell colors), was completed in less than one month, at about one-eighth of the cost estimated to do the job with a multiplying punch, whereas two or three years would have been required to complete the work using ordinary methods of calculation.

Various methods have been used to compare effects of different diets on animals. Some research workers have compared average growth of experimental and control animals over a definite period. Others have merely visually compared growth curves, etc. A more precise method of making use of all the weekly weighings is to use some algebraic expression for the growth curve, and compare the growth of the animals on various diets by comparing the coefficients of their growth equations. Various types of equations were tried by the author on this type of data but orthogonal polynomials seemed to fit the actual data better, and moreover one can add or delete a higher power without recalculation of the coefficients of preceding powers. While this study was being developed in the Food and Drug Administration, Wishart published a paper[18] wherein he fitted orthogonal polynomials to the individual weekly weights of pigs over a period of 17 weeks. He studied the effects of the diets by applying the analysis of variance technique to the orthogonal polynomial coefficients, a technique similar to that used on coefficients of Fourier series fitted to economic data.

Instead of fitting the polynomial

$$y = a + bx + cx^2 + dx^3 + \cdots$$

to a series of observations at equal intervals (where y is the weight of the rat at any week number in our present problem and x is the week number) Fisher[19] proposed a modification of Tchebycheff's method by using an equivalent polynomial

$$Y = A + B\xi_1 + C\xi_2 + D\xi_3 + \cdots$$

where ξ_1, ξ_2, ξ_3, $\cdots$ are orthogonal polynomials of degrees 1, 2, 3, $\cdots$ and where $A = \bar{y}$ and $\xi_1 = x - \bar{x}$, $\xi_2 = a'x^2 + b'x + c'$, etc.

The coefficients of the orthogonal polynomials are found by calculating the sums of the products of the y values and the ξ' values (where y equals the observed weight for any one week) and dividing by the sum of squares of the ξ' values (which are given in Fisher and Yates tables[20]).

[18] J. Wishart, "Statistical Treatment of Animal Experiments," *Supplement to the Journal of the Royal Statistical Society*, Vol. VI, No.1, 1939.

[19] R. A. Fisher, *Statistical Methods for Research Workers*, Oliver & Boyd, Ed. 5, p. 139 (1934).

[20] Table XXIII of *Statistical Tables for Biological, Agricultural, and Medical Research*, by R. A. Fisher, and F. Yates, Oliver & Boyd (1938), gives numerical values for ξ' (where $\xi' = \lambda\xi$) for all ξ' from

The coefficients are given by the algebraic expressions:

$$B' = \frac{S(y\xi_1')}{S(\xi_1'^2)} \qquad C' = \frac{S(y\xi_2')}{S(\xi_2'^2)} \quad \text{etc.}$$

Since the ξ''s are merely multiples of the ξ values, the $A, B, C, \cdots$ may be easily obtained from the $A', B', C', \cdots$ values.

In order to adapt this method to Hollerith machine punch cards, it would be necessary first to code the values given in the Fisher and Yates tables by the addition of some power of ten so as to avoid negative numbers. This can be illustrated by giving part of the table of ξ' values for $n'=52$, as given in the Fisher and Yates tables, page 60, together with the ξ'' which are the coded values of the ξ' (see Tables I and II). The values of A, B, C, etc., can be obtained by using the relationships

$$A = \frac{S(y)}{52}$$

$$B\xi_1 = \frac{S(\xi_1'Y)}{46{,}852}\,\xi_1' = \frac{2S(\xi_1'y)}{46{,}852}\,\xi_1 = \frac{S(y\xi_1'') - 100Sy}{23{,}426}\,\xi_1$$

$$C\xi_2 = \frac{S(y\xi_2'') - 1000Sy}{2{,}108{,}340}\,\xi_2$$

$$D\xi_3 = \frac{S(y\xi_3'') - 10{,}000Sy}{243{,}513{,}270}\,\xi_3$$

$$E\xi_4 = \frac{S(y\xi_4'') - 10{,}000Sy}{2{,}597{,}474{,}880}\,\xi_4$$

$$F\xi_5 = \frac{S(y\xi_5'') - 100{,}000Sy}{527{,}169{,}333{,}600}\,\xi_5.$$

The quantity $S(y\xi''_{1+2+\cdots+5})$ may also be calculated and the identity $S(y\xi''_{1+2+\cdots+5}) = S(\xi_1''y) + \cdots + S(\xi_5''y)$ used to check the results. As a further check, the relationship

$$Z = B + 90C + 10{,}395D + 110{,}880E + 22{,}503{,}600F$$

may be employed where

$$Z = \frac{S(y\xi''_{1+2+\cdots+5}) - 121{,}100Sy}{23{,}426}.$$

ξ_1' to ξ_5', and for all n' from 3 to 52. R. L. Anderson and E. E. Houseman of Iowa State College at Ames, Iowa, have recently published an auxiliary table of ξ' values for larger values of n', "Tables of Orthogonal Polynomial Values extended to $N=104$," Iowa State Agricultural Experiment Station Statistical Section, *Research Bulletin* No. 297.

TABLE I

TABLE OF ξ' VALUES AS GIVEN BY FISHER AND YATES

Obs. No.	ξ_1'	ξ_2'	ξ_3'	ξ_4'	ξ_5'
1	−51	425	−4 165	3 570	−55 930
2	−49	375	−3 185	2 170	−23 030
3	−47	327	−2 303	1 022	658
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
51	49	375	3 185	2 170	23 030
52	51	425	4 165	3 570	55 930

TABLE II

TABLE OF CODED ξ' VALUES OR ξ'' VALUES

Obs. No.	ξ_1''	ξ_2''	ξ_3''	ξ_4''	ξ_5''	$\xi''_{1+2+\cdots+5}$*
01	049	1 425	05 835	13 570	044 070	064 949
02	051	1 375	06 815	12 170	076 970	097 381
03	053	1 327	07 697	11 022	100 658	120 757
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
51	149	1 375	13 185	12 170	123 030	149 909
52	151	1 425	14 165	13 570	155 930	185 241

* Check column $\xi''_{1+2+\cdots+5} = \xi_1'' + \xi_2'' + \cdots + \xi_5''$.

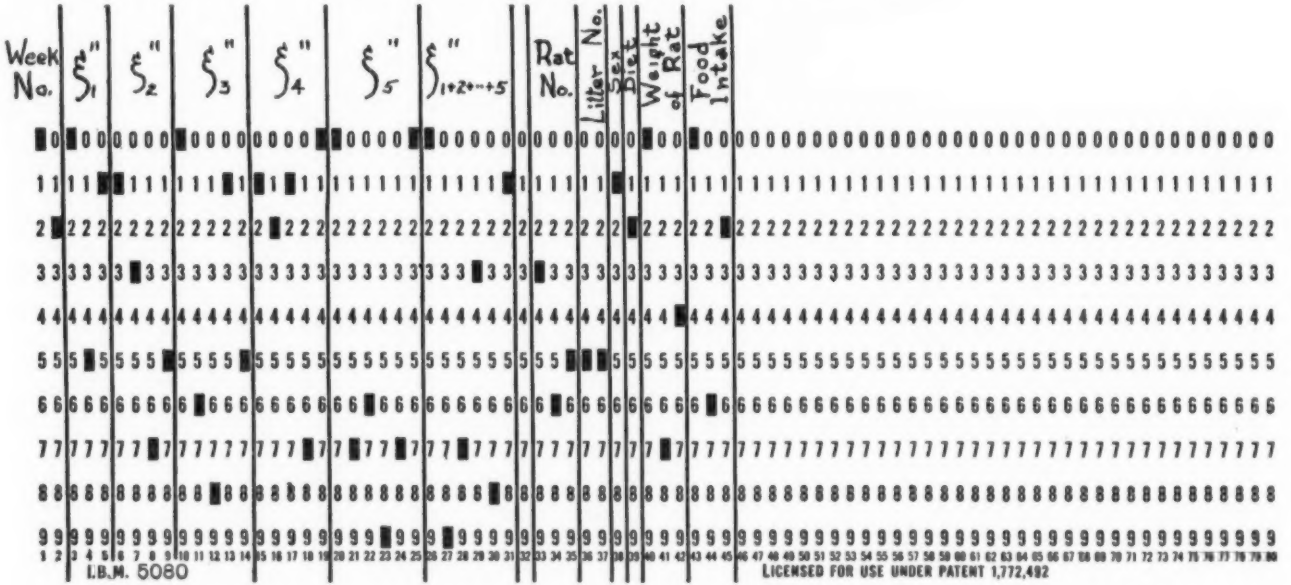
The use of punch cards facilitates the calculation of the Sy, $S(y\xi_1'')$, $S(y\xi_2'') \cdots$. One set of 52 cards is punched with the observation (or week) number, the ξ_1'', ξ_2'', $\xi_3'' \cdots$ values, and the check value $\xi_{1+2+\cdots+5}$. Referring again to Table II, one card is made for each *row* (not column). By means of a gang punch as many sets (of 52 cards each) are duplicated as there are curves to be fitted. Then taking one set of cards, it is a simple matter to punch in each card the rat or animal identification number, the litter number, the weight in grams of the animal, and its food consumption in grams for that particular week.

In order to illustrate more clearly the method used, an example will be given using data from an experiment conducted in the Food and Drug Administration.[21] In this experiment four different diets were fed to rats for the purpose of ascertaining the effect of the diet on the growth of the rat. Four rats were selected from a litter at weaning and one rat placed on each of the diets at random. Individual ad lib feedings were used. Chart I shows a card punched for week number 2 for rat number 365. (Incidentally, printed card forms are not necessary, plain cards may be used and a punched card code kept on file.)

[21] Data used in the description of this method were obtained from the Division of Pharmacology (Dr. Herbert O. Calvery, Chief) in the Food and Drug Administration where a large number of chronic toxicity studies are being made over long periods of time. The details of the experimental findings will be published elsewhere.

CHART I

PUNCHED CARD FOR WEEK NUMBER 2 OF RAT NUMBER 365



The sums of cross products can be obtained mechanically in the following manner. The punch cards for all rats are sorted (using the Hollerith sorter) on the units column of the weight and removed from the sorter in reverse order (i.e., the 9's first, then the 8's, etc.), then sorted on the rat number. Having the tabulator set for progressive totals, tabulations are made on each of the ξ'' values, as illustrated for rat number 365 by the first part of Table III, clearing the dials of the

TABLE III

Wt. of Rat Units Col.	Rat No.	Card Count	S_1	S_2	S_3	S_4	S_5	$S_{1+2\ldots5}$
9	365	5*	000313	0005781	00047671	00051103	000461971	000566839
8	365	10*	000712	0010978	00098288	00097396	000968456	001175830
7	365	17*	001423	0018245	00165009	00158429	001708413	002051519
6	365	22*	001966	0022730	00217794	00214590	002200050	002657130
5	365	26*	002362	0025856	00258332	00260456	002589362	003136368
4	365	33*	003155	0033723	00331247	00333550	003184046	003885721
3	365	39*	003841	0039249	00387561	00392933	003811965	004635549
2	365	46*	004624	0046502	00458772	00460336	004614858	005585092
1	365	48*	004766	0048522	00481092	00478611	004823007	005835998
0	365	52	005200	0052000	00520000	00520000	005200000	006297200
Tens Col.								
90	365	2*	001640	0017120	00238820	00198920	001516640	001973140
80	365	4*	003000	0038140	00457700	00379140	003671600	004549580
70	365	7*	004990	0070770	00766670	00675780	006311360	007829510
60	365	10*	006960	0103140	01094400	00936050	009720270	011860820
50	365	12*	008180	0126420	01293880	01118590	012065290	014612360
40	365	14*	009320	0151480	01463780	01335810	013746230	016706620
30	365	18*	014420	0190780	01804940	01695670	017668610	021374420
20	365	32*	031720	0332660	03075740	02992980	031814600	038247700
10	365	43*	043190	0441870	04188930	04199050	043978290	052851330
00	365	52	052000	0520000	05200000	05200000	052000000	062972000
Hundreds Col.								
300	365	35*	409100	3325300	34482300	35865500	346467900	420550100
200	365	46*	487600	4416000	47219000	45489400	460460000	558072000
100	365	50*	510000	4920000	50735000	49426000	507896000	613487000
000	365	52	520000	5200000	52000000	52000000	520000000	629720000
			Totals of Summary Cards					
	365		1553282	14385266	149266926	146760294	1479678858	1791644626

Orthogonal Polynomial Coefficients for Rat #365

$B = 3.2734$ $C = -.18058$ $D = .006599$ $E = -.0003464$ $F = .00000584$ $S(y) = 14766$ and $A = 283.962$

NOTE: Newer machines equipped with list rolling devices can produce the results shown without the use of summary cards, still further speeding up the process.

* Indicates summary cards used in the totals below.

tabulator after each rat number. By means of a summary punch, the rat number, the units column of the weight, and the progressive totals of the ξ'' columns are punched on summary cards as they are tabulated, one summary card for each row in Table III. The summary cards for the zeros on the units column of the weight are sorted out after comparing to see that they are identical for all rats—the figure in each column

on these "zero cards" should be 52 times the coding factor (which is in this case some power of 10). These punch cards are then discarded. These "zero cards" were checked and sorted out in the next two sets of progressive totals also. As indicated in Table III only the starred values are used in the totals at the bottom of the table. Caution must be exercised to insure a complete series of numbers for the progressive digiting, though the series may begin at, say, 3 or 300 instead of 9 or 900. If some number is missing in the progression, an extra *summary* card should be punched which is identical with the summary card preceding the omission and this extra summary card included with the other summary punch cards used to tabulate the totals of the summary cards.

The original punch cards are again sorted on the *tens* column of the weight and then on the rat number and tabulated the same as for the units column sort. Summary cards are also punched as before, however changing the wiring in the plug board so that a zero is automatically added after each total as shown in the second part of Table III headed "tens column." The original punch cards are also sorted on the hundreds column of the weight and the same procedure followed as for the previous groups, this time two zeros are added automatically after each total on the summary cards as shown in the third section of Table III headed "hundreds column."

All the summary cards, not including the 3 groups of "zero cards" that were previously sorted out and checked, are then sorted on rat number and tabulated controlling on the rat number to obtain the totals of the summary cards for each rat. This gives the $S(y\xi_1'')$, $S(y\xi_2'')$, $S(y\xi_3'')$ $\cdots$ which were used in the formulae for obtaining the B, C, D, $\cdots$ values. The entire process is illustrated in Table III. Summary cards used in the total are indicated in Table III by asterisks. The last line of the table indicates totals of summary punch cards or the $S(y\xi_i'')$. The original punch cards are also tabulated controlling on rat number to obtain total weight (Sy) and total food consumption. The method of progressive digiting which is a mechanical summation to obtain sums of squares and cross products is explained by Dr. Brandt.[22] Time can be saved using this punched card method even if only 10 or 15 curves are to be fitted.

As Wishart[23] points out, there is a biological interpretation to the coefficients of the orthogonal polynomials. The "A" coefficient for any one curve is the average weight of the rat over the 52 week period, the "B" coefficient is the average gain in weight per week, the "C" coeffi-

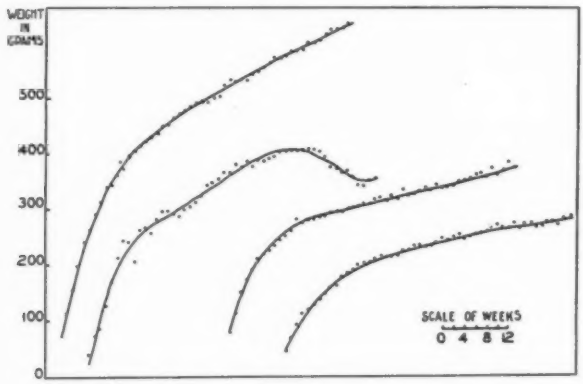
[22] *Practical Applications of the Punched Card Method in Colleges and Universities*, edited by G. W. Bahne, Columbia University Press (1935), Part X, pp. 423–436, entitled "Uses of the Progressive Digit Method," by A. E. Brandt.

[23] See footnote 18.

cient is proportional to the change in gain per week, etc. R. A. Fisher states in discussing Wishart's paper that the curve obtained using the averages of the coefficients of the orthogonal polynomials fitted to the individual animals is in fact identical with that obtained by a least squares fit to the weekly averages. The logarithms of the weights may

CHART II

FIFTH DEGREE POLYNOMIALS FITTED TO WEEKLY WEIGHINGS FOR ONE YEAR OF FOUR RATS SHOWING DIFFERENT DEGREES OF GROWTH VARIABILITY



be used instead of the weights; in fact, Wishart seems to prefer logarithms for the data used in his study.

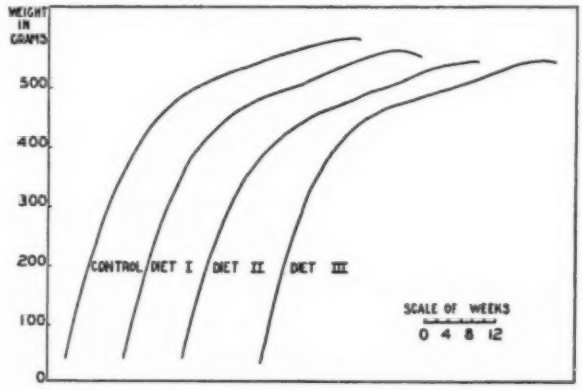
The question was raised as to the validity of the usual significance tests because of the interdependence of the observations. Stated another way, it was questioned whether the highly interdependent weekly weighings on one rat should be used to determine the degree of the curve fitted since the usual significance tests are based on an assumption of independence of the data. (Other aspects of this same question of serial correlation have been discussed by many writers. A partial list of references was given in an article in this JOURNAL[24] in 1940.) It may be argued that the interdependence is of a negligible amount in growth curves of rats. To investigate this question individual weighings on 100 comparable rats fed the same control diet over a period of one year were assembled. Using this data, instead of calculating for each rat the serial correlation coefficient as it is usually defined (i.e., the correlation between successive observations on the same animal) a slightly different type of correlation was employed, namely, the correlation between the weight of a rat one week and its weight some one following week, with

[24] L. F. Knudsen, *Interdependence in a Series*, this JOURNAL, Vol. 35, p. 507 (1940).

$N = 100$. In this way a more definite picture could be obtained of the estimated amount of interdependence at various points on a general growth curve, after the average "trend" has been removed. These correlation coefficients varied from .47 between the weight of the rat the 3rd week and its weight the 26th week to .998 between the weights for

CHART III

AVERAGE GROWTH CURVE FOR EIGHT MALE RATS ON EACH DIET OVER A PERIOD OF 52 WEEKS



the 25th and 26th weeks. In no case was the correlation between adjacent weeks less than .9.

In the study of the Food and Drug Administration data which has been used as an example, fifth degree polynomials seemed to fit the data better than a polynomial of lower degree and as well as a polynomial of higher degree. Chart II shows the weekly weights of four rats for one year and the fifth degree polynomials fitted to them. The second and third curves in Chart II show the type of difference that may exist in growths of animals, while the overall gain in weight over that period may be identical. Gain in weight (the same as a first difference) assumes a straight line regression of weight with time. It can be seen from inspection of Chart II that a *considerable amount of serial correlation exists even in the residuals* from this curve, i.e., an observation above the curve is usually followed by one or more successive observations above the curve. The significance of the coefficients might be tested by the usual method using an error based on the sum of squares of residuals from a single curve. However, because of the high degree of interdependence between the residuals from the curve, it seemed desirable to use a test of the average coefficient based on the scatter of the coefficients of

several rats of the same diet and sex about that average. The averages and standard deviations of the individual coefficients (the A's, the B's, the C's, etc., separately) were calculated for each diet and sex. The standard error of the average A, B, or C, etc., coefficient was calculated in the usual way, that is, by dividing the standard deviation by the square root of N. In other words, the standard error of each average coefficient was calculated by considering the scatter of the *individual coefficients about the average*. As a test of significance of the coefficients, Student's "t" test was applied to test whether the average coefficient differed significantly from zero. This procedure avoids any assumption of independence of the weekly weights. Furthermore, the A, B, C, $\cdots$ coefficients are independently distributed.

Table IV gives an idea of the size of the average coefficients in the example used and their standard errors; and Chart III shows the growth curves described by these coefficients. The average coefficients includ-

TABLE IV

AVERAGE ORTHOGONAL POLYNOMIAL COEFFICIENTS FOR THE CONTROL AND THE THREE EXPERIMENTAL DIETS

Male 8 Rats	A $\text{av} \pm \sigma_{\text{av}}$*	B $\text{av} \pm \sigma_{\text{av}}$*	C $\text{av} \pm \sigma_{\text{av}}$*	D $\text{av} \pm \sigma_{\text{av}}$*	E $\text{av} \pm \sigma_{\text{av}}$*
Control	461.77 ± 21.38	8.211 ± .809	−.2995 ± .0303	.00960 ± .00086	−.000273 ± .000037
Diet I	443.22 ± 20.58	7.673 ± .479	−.2759 ± .0214	.00889 ± .00058	−.000373 ± .000024
Diet II	422.54 ± 25.10	7.531 ± .699	−.2517 ± .0144	.00907 ± .00119	−.000299 ± .000029
Diet III	438.07 ± 19.74	7.312 ± .540	−.2698 ± .0188	.01306 ± .00085	−.000390 ± .000040

* The average of the indicated coefficient for all the rats of the same sex on each diet ± the standard error of the average.

NOTE: This table is given merely for the purpose of illustrating the typical results to be expected and the size of the errors.

ing those of fourth and fifth degree were significantly different from zero, showing that a polynomial of not less than the fifth degree should be used in this problem. As contrasted with Wishart's study, where he uses a third degree polynomial for a 17-week growth period, it seems logical that a 52-week growth period would require a slightly more complicated curve.

The effects of the diets on the growth of the rat can be tested by using analysis of variance either on the coefficients themselves or on the coefficients corrected for the influence of initial weight and/or the food intake.

In growth studies there is good reason to expect a high degree of interdependence between residuals from growth curves. Consequently, instead of employing the usual tests of significance it may be advisable to estimate the standard error of the coefficients from the scatter of the individual coefficients about their average.

A NEW MULTIPLICATIVE SEASONAL INDEX*

By Julius Shiskin
National Bureau of Economic Research

This paper proposes a new multiplicative seasonal index, describes a quick way to compute it, and compares the new index with the ratio-to-moving-average index and with the monthly means index. The new index has some merits of its own, but the chief value of this paper lies in the light it throws on the relations among these indexes.

Derivation and computation. Let us consider a monthly time series for which a scatter diagram is drawn so that values for a given month (e.g., the January values) are plotted as the ordinate and the corresponding values representing the trend and cyclical components (hereafter referred to as the trend-cycle component) as the abscissa. If the original values for the month include neither a random nor a seasonal component, all the points fall on a straight line that passes through the origin and has a slope of one, since the trend-cycle component has merely been plotted against itself along axes with identical scales. If the assumptions are changed to allow a multiplicative seasonal component in the original values, all the points fall on a straight line that passes through the origin, but the slope deviates from one. If the original values include an additive seasonal component, the slope of the line remains one, but the line no longer passes through the origin. If the seasonal component is partly additive and partly multiplicative, the line does not pass through the origin and its slope differs from one. These relations tend to prevail if the series also includes a random component. However, the observations no longer fall on a straight line, but tend to be distributed at random around such a line. It can be concluded, therefore, that the seasonal component for a given month can be measured by the difference between the parameters of a fitted straight line and the parameters of a line passing through the origin and having a slope of one.[1]

The relations between the seasonal and the trend-cycle components may be expressed as follows. Start with the general linear equation

* This paper summarizes a section of a study on the measurement and elimination of seasonal fluctuations, in which particular attention is paid to the general method described below.

The writer is under heavy obligation to Dr. Arthur F. Burns for many contributions to this work. Valuable suggestions were also made by Dr. Simon Kuznets, Mr. Bert Gottfried, and Mr. Denis Volkenau.

[1] This approach to the study of seasonal fluctuations was suggested to the writer by Mr. Milton Friedman and Mr. Gerhard Bry. It appears in "Ausschaltung von Saisonschwankungen mittels Lag-Correlation," by F. I. Zrzavy, *Monatsberichte des Oesterreichischen Instituts für Konjunkturforschung*, Beilage n. 2, Vienna, October 1933; and "Note on Seasonal Variation" by J. Wisniewski, *Studia Ekonomiczne*, Vol. 3, Cracow 1936. It is a simplification of a method described by Horst Mendershausen in "Eliminating Changing Seasonals by Multiple Regression Analysis," *Review of Economic Statistics* November 1939, pp. 171–177.

$y_r = a + bx$, where y_r denotes a regression value for a given month as distinguished from y, an actual value of the original series, and x denotes the corresponding trend-cycle value. As long as there is no seasonal component, $y_r = x$. If the seasonal component is multiplicative, $y_r = bx$, and $b \neq 1$. If the seasonal component[2] is additive, $y_r = a + x$.

It is instructive to examine a special case of the general linear equation, the case where $b = 1$, or is assumed equal to 1. The value of a can now be evaluated for an empirical series from the equation:

$$\sum y_r = \sum y = na + b \sum x,$$

where n is the number of years covered by the given month. Here

$$a = \frac{\sum y - \sum x}{n} = \frac{\sum (y - x)}{n}.$$

The last expression is precisely the formula used for deriving an additive seasonal index by averaging actual differences of original values from the trend-cycle curve. Hence, in this special case, a is the uncentered[3] additive seasonal index.

Another special case is where a is zero, or is assumed equal to zero. Here $\sum y = b \sum x$ and $b = \sum y / \sum x$.

$\sum y / \sum x$ is suggested as a new multiplicative seasonal index.[4] It may be calculated as follows. (1) The sums of the original figures and the sums of the trend-cycle values for each month are computed. In the computations it is convenient to represent the trend-cycle component by the twelve-month moving average of the original series.[5] (2) The

[2] It is assumed here that the random factor has no disturbing effect upon these relations. This assumption is valid in principle, but not in a particular series.

[3] That is, the indexes for the twelve months (the discussion in the text refers to any one month) will not sum to zero. The indexes are centered by subtracting from each one-twelfth the difference between the sum of the uncentered indexes and zero. One consequence of this adjustment is that for any year the sum of the original figures equals the sum of the seasonally-adjusted figures.

[4] It is important to note that the multiplicative seasonal factor is assumed here to apply to the trend-cycle component, not to the trend-cycle-random component as in the standard methods.

A simple procedure for converting the additive into the multiplicative index, and vice versa, is indicated by the expression obtained when the equation, $\Sigma y = na + b\Sigma x$, is divided through by Σx:

$$\frac{\Sigma y}{\Sigma x} = a \frac{n}{\Sigma x} + b = \frac{a}{\bar{x}} + b,$$

where $\bar{x}$ is the mean trend-cycle value for a given month.

If the additive seasonal index is already computed, b is assumed to be one, and $\Sigma y / \Sigma x$ therefore can be derived by dividing the additive index for each month by the mean trend-cycle value for that month, and then adding one. If the multiplicative index is computed, the additive index can be derived by subtracting one from the multiplicative index for each month and multiplying the result by the corresponding mean trend-cycle value. If the seasonal correction factors have already been obtained from the general expression, $y_r = a + bx$, and the multiplicative index is desired, it can be derived by substituting the values of a, b, and $\bar{x}$ in the equation, and solving for $\Sigma y / \Sigma x$.

[5] The twelve-month moving average is defective as a measure of the trend-cycle component in that it reflects changes in seasonality, it does not reach sufficiently up into cyclical peaks nor down into troughs, it does not always iron out random movements of a few months' duration, it does not cover

sum of the original figures for each month is divided by the sum of the moving-average values for that month. (3) In order that the relatives given by the preceding step sum to 1,200 (or average 100), each of the twelve relatives is expressed as a percentage ratio to the arithmetic mean of the twelve relatives.

In this procedure, it is not necessary to compute the moving average.[6] If a twelve-month moving total is calculated and centered at the seventh month,[7] as is common practice, the twelve-month moving totals for the Julies are identical with the annual totals. Let a series run from y_1 to y_{49}, where y_1 is the value for January of the first year, y_2 the value for February of the first year, and so on. The moving-average values for the successive Julies are

$$\frac{\sum_{1}^{12} y}{12}, \quad \frac{\sum_{13}^{24} y}{12}, \quad \frac{\sum_{25}^{36} y}{12}, \quad \text{and} \quad \frac{\sum_{37}^{48} y}{12},$$

and the sum of the July moving-average values is

$$\frac{\sum_{1}^{48} y}{12}.$$

The moving-average values for the successive Augusts are

$$\frac{\sum_{2}^{13} y}{12}, \quad \frac{\sum_{14}^{25} y}{12}, \quad \frac{\sum_{26}^{37} y}{12}, \quad \text{and} \quad \frac{\sum_{38}^{49} y}{12},$$

and the sum of the August moving-average values is

$$\frac{\sum_{2}^{49} y}{12}.$$

But the sum of the moving-total values for the Augusts can be derived from that for the Julies by subtracting y_1 and adding y_{49}, and the sums for other months can be obtained by similar calculations.[8]

---

the full series, and it may turn as much as six months before or after the original series. These defects usually have little influence on seasonal indexes.

[6] I am indebted to Mr. Bert Gottfried for pointing out this simplification.

[7] The effect upon the seasonal index of centering the twelve-month moving average on the seventh month instead of between the sixth and seventh months is negligible.

[8] This short cut can, of course, be used to find the mean moving-average values in other problems; for example—in computing the difference-from-moving-average index by the formula $\frac{\Sigma y - \Sigma x}{n}$, or in transforming an additive index to a multiplicative index.

Relation to the ratio-to-moving-average index. There is an enlightening relation between $\sum y/\sum x$ and the ordinary unweighted mean ratio-to-moving-average index, $\sum(y/x)/n$. The former is a weighted mean ratio-to-moving-average index, where the weights are the moving-average values. This is plain from the equation below, in which a single prime denotes the first year, a double prime the second year, and so on.

$$\frac{\frac{y'}{x'}\cdot x' + \frac{y''}{x''}\cdot x'' + \frac{y'''}{x'''}\cdot x''' + \cdots}{\sum x} = \frac{\sum y}{\sum x}.$$

Since the new index differs from the ordinary ratio-to-moving-average index only in that it is weighted by the moving-average values, it seems appropriate to call it the weighted ratio-to-moving-average index, or more briefly, the weighted-ratio index.

As is well known the relation between a weighted and an unweighted mean depends upon the correlation between the variates and the weights. If the Pearsonian coefficient of correlation is positive, the weighted mean will be greater than the unweighted mean; if it is negative, the weighted mean will be smaller; and if the correlation is zero, the two means will be equal. Hence, when there is little correlation between y/x and x, $\sum y/\sum x$ will tend to be a good approximation of $\sum(y/x)/n$.

A further difference between the indexes arises from the way they are centered, since the weighted-ratio index is centered by dividing each preliminary index by the average of the preliminary weighted-ratio indexes, and the ratio index is centered by dividing each preliminary index by the average of the preliminary ratio indexes.

For a given month differences in ratios are assumed to arise from random causes. In computing a seasonal index, a given random factor should receive the same weight wherever it falls. Under the weighted-ratio method, a random factor which falls at a high level receives a greater weight than a similar random factor which falls at a low level. Since the ratio method gives the same weight to all ratios, it seems superior to the weighted-ratio method in this respect.

Another advantage of the ratio method is that the ratios computed for each month can be put to a useful purpose. A chart of the ratios for successive Januaries, Februaries, and so on, is helpful in determining whether a changing or a constant seasonal pattern prevails, and what periods are appropriate in the latter case. The weighted-ratio method assumes a constant pattern and relies upon a study of the original series for the selection of periods for which this assumption seems reasonable.

A further advantage of the ratio method is that extreme items can easily be excluded from the indexes. They can be picked from the ratio charts for individual months and omitted, or one of several positional means can be selected. The elimination of extreme items is more difficult if the weighted-ratio method is used, for extreme items do not show up so clearly in charts of the original data as in charts of ratios to moving averages, and positional means based on arrays of original figures are highly suspect, particularly in series that have pronounced trends or cycles.

Relation to the monthly means index. The new index bears a close resemblance to the familiar monthly means index. This index is constructed as follows. (1) Totals of the original figures for each month are computed. (2) These totals are expressed as relatives of the arithmetic mean of the twelve totals. (3) The relatives are adjusted for trend. The trend correction can be made by computing the slope of a least-squares straight line fitted to annual averages of the monthly data expressed as relatives of their arithmetic mean and applying this slope to the relatives. The calculations are simplified by taking the origin at the middle year. The slope can then be calculated directly from the formula $b=\sum xy/\sum x^2$, where x refers to the year and y to the annual averages in relative form. Since b gives the annual increment, it must be divided by 12 to yield the monthly increment. If July is assumed to be the central point, no adjustment of the relative for this month need be made. The other months can then be corrected for trend by subtracting b' (where $b'=b/12$) from August, $2b'$ from September, $3b'$ from October, $4b'$ from November, and $5b'$ from December; and by adding b' to June, $2b'$ to May, $3b'$ to April, $4b'$ to March, $5b'$ to February, and $6b'$ to January.

The relations between the monthly means index and the weighted-ratio index may be set out as follows. Let $A, B, C, \cdots L$ refer to the months of the year, and the subscripts $1, 2, 3, \cdots n$ to the years, so that A_1 is the value for January of the first year, G_3 the value for July of the third year, L_5 the value for December of the fifth year, and so on. Assume that the moving average is available for the full period.

The uncentered weighted-ratio index for July is

$$\frac{12\sum G}{A_1 + B_1 + C_1 + \cdots + L_n}.$$

Since the monthly means method, as described here, does not require a trend correction for July, this is also the formula for the monthly means index for July.

The uncentered weighted-ratio index for August is

$$\frac{12\sum H}{B_1 + C_1 + D_1 + \cdots + A_{n+1}}.$$

The monthly means index for August is

$$\frac{12\sum H}{A_1 + B_1 + C_1 + \cdots + L_n},$$

with a trend correction applied.

The difference between the two August indexes is that the monthly means index includes A_1 and not A_{n+1}, and is explicitly corrected for trend. Similar differences characterize other months. An additional, though minor, difference arises when the weighted-ratio index is centered.

The monthly means index for July may be written:

$$12\left[\frac{G_1}{A_1 + B_1 + \cdots + L_1}(A_1 + B_1 + \cdots + L_1) + \frac{G_2}{A_2 + B_2 + \cdots + L_2}(A_2 + B_2 + \cdots + L_2) + \cdots\right] \div (A_1 + B_1 + \cdots + L_n).$$

That is, the monthly means index may also be written as a weighted average of ratios. For this reason, it may be said to have the same kind of implicit weighting as the weighted-ratio index.

The weighted-ratio method assumes that the random component and most of the cyclical component in the series cancel out when the original figures for any month are summed, and the trend and the residual cycle are eliminated when the sum of the original figures is divided by the sum of the moving-average values. The latter assumption is valid when the curve used to measure the trend-cycle component is a perfect measure of it. If the twelve-month-moving-average is used in the manner suggested here, the trend-cycle adjustment is crude, for the difference between the trend adjustments for two successive months is nothing more than the difference between the figure from the end of the series which is added and the figure from the beginning of the series which is dropped, e.g., $A_{n+1} - A_1$. This difference may be dominated by random factors. The monthly means method utilizes the slope of a straight line fitted to annual totals to adjust for the trend and residual cycle. This method has the advantage that the trend adjustment is

more stable from month to month, but it suffers from all the difficulties connected with fitting a trend, and in addition from the fact that it adjusts only for a linear trend. On theoretical grounds alone it is difficult to say that one of these trend adjustments is better than the other.[9]

A disadvantage of the new method is that it involves the loss of a half year at each end of the series. The monthly means method must be confined to complete years, but this does not involve so great a loss, since seasonal indexes can be computed on a fiscal-year basis. This disadvantage of the weighted-ratio index is not very serious, partly because most series are divided into several seasonal periods, and only the first and last periods are affected, and partly because the moving average can be made to cover the full period by free-hand extrapolation.

Comparisons of indexes for particular series. Seasonal indexes have been computed by the weighted-ratio method, the monthly means method, and the ratio method for ethyl alcohol production in the United States, 1921–40, refined copper shipments in the United States, 1927–36, and Douglas fir production in the United States, 1918–27, and time records have been kept. Table I shows that the weighted-ratio method

TABLE I

TIME REQUIRED TO COMPUTE MULTIPLICATIVE SEASONAL INDEXES BY DIFFERENT METHODS

Method	Ethyl alcohol production, 1921–40		Refined copper shipments, 1927–36		Douglas fir production, 1918–27	
	Minutes	Per cent of time for weighted-ratio method	Minutes	Per cent of time for weighted-ratio method	Minutes	Per cent of time for weighted-ratio method
Monthly means	57	119	42	108	34	110
Ratio	181	377	97	249	95	306
Weighted ratio	48	100	39	100	31	100

took the least time in each case. In this respect its advantage over the ratio method is considerable, and its advantage over the monthly means method small but persistent. It must be pointed out, however, that the time advantage over the monthly means method is not conclusive, since the monthly means method will be somewhat shorter if the linear trend

[9] It is worth pointing out that the difference-from-twelve-month-moving-average index is not, as is nearly universally assumed, obviously superior to the additive monthly means index. The difference-from-moving-average index for a given month is $\frac{\Sigma(y-x)}{n}$, or $\frac{\Sigma y-\Sigma x}{n}$ and the term which adjusts for trend is Σx. Since this is precisely the term which adjusts for trend in the weighted-ratio index, the difference-from-moving-average index, like the weighted-ratio index (and unlike the ratio index), involves a crude trend correction.

is derived by the method of selected points instead of least squares; though this procedure may yield a cruder index.

Seasonal indexes have also been computed by each of the three methods for seven additional series.[10] The centered indexes for the ten series are given in Table III. Table II compares the indexes by the device of two summary measures. The first, the coefficient of similarity, is computed as follows. (1) The deviations from 100 of each of the pair of seasonal indexes under comparison are computed. (2) One set of deviations is adjusted so that its sum is equal to the sum of the second set, which is considered the base. (3) The differences between the adjusted deviations, month for month, are summed without regard to sign. (4) Finally the sum of the differences is divided by the base figure and one subtracted from this quotient.[11] Since this measure eliminates differences in amplitude, the table gives another measure which

TABLE II

COMPARISON OF SEASONAL INDEXES COMPUTED BY DIFFERENT METHODS

Series	Coefficient of similarity			Average difference		
	Weighted-ratio and ratio index	Monthly means and ratio index	Monthly means and weighted-ratio index	Weighted-ratio and ratio index	Monthly means and ratio index	Monthly means and weighted-ratio index
1. Steel ingot production in U.S., 1921–38	.892	.895	.956	0.54	0.58	0.35
2. Refined copper shipments in U.S., 1927–36	.887	.549	.509	0.42	1.64	1.77
3. Douglas fir production in U.S., 1918–27	.933	.912	.944	0.53	0.68	0.35
4. Ethyl alcohol production in U.S., 1921–40	.969	.970	.972	0.78	0.37	0.72
5. Crude cottonseed oil production in U.S., 1918–24	.977	.985	.970	1.53	1.06	1.82
6. Cotton consumption in U.S., 1914–22	.926	.807	.858	0.18	0.59	0.47
7. Cotton consumption in U.S., 1923–39	.966	.944	.948	0.29	0.51	0.28
8. Corporate issues of stocks in U.S., 1907–27	.699	.655	.921	4.22	5.05	1.25
9. New corporate issues of bonds and stocks in France, 1922–38	.584	.590	.955	12.52	12.56	1.26
10. London bank clearings, 1868–1913	.934	.902	.945	0.24	0.32	0.22

[10] Strictly comparable procedures have been used in computing all ten indexes. Periods for which the moving average is fully available were selected and no attempt was made to omit extreme items.

[11] See Simon Kuznets, *Seasonal Variations in Industry and Trade* (National Bureau of Economic Research, 1933) pp. 281–282.

TABLE III

SEASONAL INDEXES COMPUTED BY THE MONTHLY MEANS METHOD, RATIO-TO-MOVING-AVERAGE METHOD, AND THE WEIGHTED RATIO-TO-MOVING-AVERAGE METHOD

Series	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Total
1. Steel ingot production in U.S., 1921–38													
Monthly means method	99.2	106.0	109.7	108.0	105.6	98.2	94.3	97.6	97.1	99.1	94.6	90.8	1200.2
Ratio method	99.3	106.6	109.1	108.2	106.6	99.2	94.2	96.6	96.4	99.4	94.8	89.7	1200.1
Weighted-ratio method	99.1	106.3	110.0	108.4	106.1	98.4	94.5	97.6	96.9	98.7	93.9	89.9	1199.8
2. Refined copper shipments in U.S., 1927–36													
Monthly means method	95.9	95.2	101.9	97.3	98.2	102.4	98.1	102.4	102.1	111.9	101.2	93.5	1200.1
Ratio method	99.0	97.8	103.4	97.7	98.9	103.8	98.1	101.8	100.8	110.5	98.4	89.6	1199.8
Weighted-ratio method	98.7	97.7	104.0	98.9	99.2	102.8	98.0	101.7	100.7	109.7	98.3	90.0	1199.7
3. Douglas fir production in U.S., 1918–27													
Monthly means method	87.2	98.9	101.9	101.7	105.9	107.6	91.5	105.0	104.7	108.0	102.6	84.9	1199.9
Ratio method	86.0	97.6	100.7	101.5	106.3	108.1	91.8	105.8	105.6	108.4	103.1	85.4	1200.3
Weighted-ratio method	86.7	98.7	101.6	101.2	105.6	107.4	91.5	105.3	105.0	108.3	103.0	85.8	1200.1
4. Ethyl alcohol production in U S., 1921–40													
Monthly means method	93.8	83.1	87.1	84.6	90.0	91.0	97.6	104.9	109.0	126.4	118.5	114.1	1200.1
Ratio method	94.9	82.9	86.4	84.7	90.2	91.0	97.1	104.4	109.4	126.7	118.6	113.8	1200.1
Weighted-ratio method	94.9	83.7	87.7	85.2	90.6	91.5	97.9	104.8	108.6	125.5	117.3	112.4	1200.1
5. Crude cottonseed oil production in U.S., 1918–24													
Monthly means method	161.0	129.1	103.8	62.4	39.1	21.3	11.9	15.3	80.8	199.4	206.1	169.8	1200.0
Ratio method	163.8	131.0	104.0	61.4	38.0	20.4	11.4	15.0	80.9	200.8	206.0	167.4	1200.1
Weighted-ratio method	164.4	132.0	106.1	63.7	39.8	21.5	11.6	14.5	79.0	196.6	203.5	167.2	1199.9
6. Cotton consumption in U.S., 1914–22													
Monthly means method	103.7	95.7	105.5	100.2	103.7	102.1	97.6	99.1	96.8	99.7	99.3	96.7	1200.1
Ratio method	104.2	96.6	106.0	100.8	104.0	102.5	97.8	98.8	96.6	99.0	98.1	95.4	1199.8
Weighted-ratio method	104.4	96.4	106.0	100.7	104.0	102.2	97.6	98.9	96.5	99.1	98.5	95.8	1200.1
7. Cotton consumption in U.S., 1923–39													
Monthly means method	106.6	99.7	109.1	102.1	101.7	95.5	88.9	95.1	96.7	105.9	102.7	96.0	1200.0
Ratio method	107.5	100.6	109.8	102.4	101.6	94.9	88.4	94.4	96.4	106.2	102.4	95.5	1200.1
Weighted-ratio method	107.1	100.2	109.5	102.3	101.7	95.5	88.8	94.8	96.4	105.6	102.3	95.6	1199.8
8. Corporate issues of stocks in U.S., 1907–27													
Monthly means method	110.7	113.0	107.4	99.7	84.4	108.0	82.0	85.1	73.7	95.3	92.4	148.3	1200.0
Ratio method	111.0	116.6	104.3	105.8	86.0	115.5	87.6	90.6	66.6	86.2	92.5	137.3	1200.0
Weighted-ratio method	111.0	113.7	108.4	100.7	85.0	109.0	83.5	86.0	74.2	94.8	91.4	142.3	1200.0
9. New corporate issues of bonds and stocks in France, 1922–38													
Monthly means method	110.0	168.7	82.5	93.3	99.5	139.5	70.3	39.4	61.6	148.3	92.2	94.9	1200.2
Ratio method	108.6	132.7	77.6	110.2	119.3	118.6	78.0	44.5	67.3	139.2	112.2	91.7	1199.9
Weighted-ratio method	109.4	165.3	81.6	91.7	99.6	139.1	70.2	38.8	62.7	152.3	93.5	95.9	1200.1
10. London bank clearings, 1868–1913													
Monthly means method	104.3	102.9	101.1	105.9	100.4	98.7	102.6	94.5	90.3	99.4	97.1	102.9	1200.1
Ratio method	104.5	103.3	101.8	105.9	100.4	98.8	102.7	95.0	90.3	98.8	96.7	102.0	1200.2
Weighted-ratio method	104.7	103.1	101.4	106.2	100.5	98.8	102.6	94.5	90.1	99.1	96.9	102.4	1200.3

takes amplitude into account—the average difference disregarding sign. The average difference is computed by subtracting one index from another, month for month, summing without regard to sign, and dividing by 12.

Tables II and III support the conclusion that there is a greater similarity between the monthly means index and the weighted-ratio index than there is between either of these indexes and the ratio index. The coefficients of similarity between the monthly means and the weighted-ratio indexes are highest in six of the ten series and the average difference is smallest in six of the ten series. The greater resemblance is most notable in the two series which contain a large random component: corporate issues of stocks in the United States, 1907–27, and new corporate issues of bonds and stocks in France, 1922–38. In such series the implicit weighting exerts its greatest effect and substantial differences between the weighted-ratio indexes and the ratio index may be expected.

The tables also suggest that the weighted-ratio index provides a better estimate of the ratio index than does the monthly means index. The new index is closer to the ratio index in eight series according to the average difference and in six series according to the coefficient of similarity. In no case is the monthly means index appreciably closer to the ratio index than is the weighted-ratio index, whereas in refined copper shipments the weighted-ratio index is much closer to the ratio index.

Summary. The weighted-ratio index has some advantages over the monthly means index. It can be computed more rapidly, it is a special case of a general method, and it can be converted to a strictly analogous additive index. The empirical evidence suggests that the weighted-ratio index gives a better approximation to the ratio index. But in the absence of strong theoretical support, the margin of superiority is too small to warrant a generalization.

Like the weighted-ratio index, the monthly means index is a centered average of ratios weighted by trend-cycle values. The two indexes almost always come out very nearly the same. They usually give close approximations to the ratio index. But if the random component or the correlation between the ratios and the corresponding trend-cycle values is large, the weighted-ratio indexes may depart considerably from the ratio index.

SEASONAL FACTORS DETERMINED BY DIFFERENCE FROM AVERAGE OF ADJACENT MONTHS

By Henry A. Latané
Lionel D. Edie and Company, Inc.

This method for smoothing economic time series is based on a comparison of the data for each individual month with the average of such data for the two adjacent months. It is assumed that, in any time series from which variations due to seasonal factors have been satisfactorily eliminated, the index for any one calendar month is typically equal to the average of the indices of the two adjacent months. For example, the adjusted February indices, over a series of years and after eliminating data obviously affected by extraneous factors such as strikes, should fluctuate around the averages of the adjacent January and March indices and typically should equal such averages. Adjustment factors which produce such results are taken to be satisfactory seasonal factors.

The method is based on the use of moving averages and so is directly comparable with other such methods now in use. Its chief advantage is that the comparison is close to the data and consequently there are fewer extraneous factors to hide the seasonal pattern. This is illustrated by showing the per cent by which railroad freight revenues[1] for each month over a series of years differs (1) from the average of the adjacent months and (2) from a twelve-month centered moving average. Examples are:

	April		July	
	(1) Per cent	(2) Per cent	(1) Per cent	(2) Per cent
1936	.4*	−3.4	3.6	1.8
1937	−2.4	−1.7	3.4	3.9
1938	−3.4	−11.0	3.6	.4
1939	−6.8*	−13.7*	3.2	−1.8
1940	−2.4	−9.6	2.6	1.0

* Comparison affected by coal strike.

When the adjacent months are used as a base for comparison the range of variation within each calendar month is consistently smaller and the effects of random events are more clearly indicated. There are seven calendar months wherein the range among the percentage differences

[1] Class I Railway Freight Revenues (I. C. C. data) adjusted to a uniform month basis but not adjusted for seasonal variation.

(from the average of adjacent months) is less than 2 percentage points, after eliminating data affected by known extraneous factors. The widest such range is nearly 7 points, but, even so, it is smaller than the range for all but two months based on the twelve-months-moving average.

The greater consistency simplifies the problem of selecting the proper amount of variation to be eliminated. The shorter base also permits use of current data earlier than is possible with other averaging techniques. Using a centered average avoids the problem of eliminating trend and thus escapes one of the major difficulties in the link-relative method. The time and skill involved in applying the average of adjacent months is approximately the same as that of other simple averaging techniques.

A corollary but important advantage of the method is that it furnishes a direct check on the work-day adjustment factor. This is of especial significance in time of war when the work week for many industries is unknown or changing. Take the work-day adjustment for railroad freight revenues as an example. The Federal Reserve Board uses a six-day week for adjusting these data. In their adjusted series for the period from July 1935 through July 1941, twenty months with five Sundays were above the average of the adjacent months, while only four were below such average. Of the "four-Sunday" months, only eleven out of forty-one were above the corresponding average. It therefore is obvious that too much allowance was made for the "five-Sunday" months. The median index for such months was 2.1 per cent above the average of the two adjacent months, indicating that adjustment for a week of $6\frac{1}{2}$ working days would be considerably more satisfactory than the six-day week actually used by the Board. The same series adjusted on a $6\frac{1}{2}$ day basis showed only 13 out of 24 "five-Sunday" months above the adjacent months average. While there is little actual origination of freight traffic on Sunday, enough is received from connections to make a $6\frac{1}{2}$ day working week entirely logical.

The average of adjacent months is valuable for checking whether or not seasonal variation has been successfully eliminated in any given series. When a test shows that the adjusted indices for any one calendar month are consistently above or below the average of adjacent months, it is a strong indication that further adjustment for seasonal is needed. If constant seasonal factors were used in the original computations, it is possible to re-compute these factors by the adjacent months method directly from the adjusted data. The adjusted Federal Reserve Board series for railroad freight revenues was reworked in this manner and new seasonal factors computed which differed very materially from those used by the Board but which showed a maximum difference of

.3 per cent from factors obtained by working directly from the original data.

The proposed method also escapes the heavy dependence on judgment which characterizes non-mathematical techniques. It is particularly valuable for determing the net difference of two series such as revenues and expenses where it is very important to keep the method for eliminating seasonal variations uniform. It was for studies of railroad net income that this method was originally developed. Here the difference between revenues and expenses is small and seasonal adjustments must be refined to be of any value. The method has been in use for this purpose for several years.

Table I shows railroad freight revenues adjusted by this method as compared with the corresponding Federal Reserve Board series. As is apparent, the differences are substantial. In another more recently adjusted Board series, used as a test,[2] seasonal factors derived by the average of adjacent months method differed by no more than .5 per cent from those of the Board.

In deriving seasonal factors by the proposed method it is necessary to determine (1) the per cent by which the data for each calendar month typically differs from the average of the adjacent months and (2) the seasonal factors necessary to eliminate this typical difference. The typical difference may be determined by averaging the months under consideration after eliminating data known to be affected by extraneous factors. The percentage variation from adjacent months[3] for each April, for example, from 1936 through 1940 has already been given. It is known that data for April 1936 and 1939 were affected by coal strikes. The average of the other three months is taken as the typical difference. None of the corresponding figures for July were affected by known extraneous factors so an average of all of these items was used. Determination of the months to be eliminated is partly a matter of judgment,

[2] By Mr. H. C. Barton, Jr., the Division of Research and Statistics, the Federal Reserve Board. Mr. Barton's comment was: "In the one experiment which we were able to make using your method on a series which had been recently adjusted (in contrast to freight revenues which we know is bad) we found that it yielded a somewhat smoother adjustment but in no case were the factors for a particular month more than .5 per cent apart." Mr. Barton was consulted as an authority on seasonal factors in economic data and his generous encouragement led to preparation of this material for publication.

[3] The arithmetical derivation of these percentage variations is simple. For example,

	Original Data	Take Difference	Take Difference	Halve	Divide by Original Data
January	90				
February	100	10	20	10	10%
March	90	−10	0	0	0
April	80	−10			

TABLE I

MONTHLY FREIGHT REVENUES, CLASS I RAILROADS
(000,000 omitted)

	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
(1) Data adjusted for 6½ day work week but not for seasonal variation												
1935	209.7	224.9	223.3	220.6	219.0	228.0	219.1	225.3	262.2	273.0	255.7	236.7
1936	240.0	257.3	246.0	254.8	261.2	266.9	282.2	277.0	300.8	313.1	313.1	297.4
1937	271.3	287.5	300.0	286.0	285.3	279.2	290.5	281.9	302.0	299.1	265.9	229.3
1938	220.5	216.0	217.3	209.6	216.2	220.7	240.4	242.7	268.9	286.0	271.8	249.1
1939	248.9	244.2	246.1	226.2	237.0	253.2	267.2	264.3	323.0	345.8	318.9	278.4
1940	281.4	270.3	260.5	263.6	278.0	284.0	298.9	298.1	332.0	334.2	325.0	312.2
1941	306.9	322.2	337.5	302.5	361.0	380.9	401.9	399.5	422.8	421.2	403.3	385.7
1942	389.1	410.8	433.9	463.8	483.6	496.8	528.3					
Proposed Seasonal Factors	95.1	94.8	96.1	93.5	95.4	96.3	102.1	100.9	110.5	114.0	105.8	95.5
(2) Data adjusted for seasonal variation by the proposed method												
1935	220	237	232	236	230	237	215	223	237	239	241	248
1936	252	272	256	272	274	277	276	274	272	275	296	311
1937	285	303	312	306	299	290	285	279	273	262	251	240
1938	232	228	226	224	227	229	235	240	243	251	257	261
1939	261	258	256	242	249	263	262	262	292	303	301	294
1940	296	285	271	282	292	295	293	295	300	293	307	327
1941	323	340	351	323	379	396	394	396	382	370	381	404
1942	409	433	452	496	507	516	518					
F R B Seasonal Factors	95	97	99	97	98	97	101	99	108	111	103	95
(3) Data adjusted for seasonal variation by Federal Reserve Board												
1935	219	232	228	226	221	238	215	226	247	244	247	252
1936	250	265	251	261	270	273	277	283	277	279	309	311
1937	289	296	300	293	294	286	285	287	278	272	257	240
1938	235	223	217	214	223	226	241	243	247	260	262	260
1939	265	251	245	236	239	259	268	265	298	315	309	297
1940	294	278	266	270	281	297	294	298	313	298	314	333
1941	321	333	345	310	365	398						

but not usually difficult. Variation in coal production is the most important single factor influencing railroad freight revenues in normal times, and the effects of the numerous bituminous and anthracite coal strikes in recent years are obvious.

The second step is to determine the seasonal factors which will eliminate the typical variation of each month from the average of adjacent months. The mathematical solution of this problem involves a set of twelve simultaneous equations. Fortunately, however, these equations can be reduced to the first degree and solved by simple cumulative addition, as follows:

Let:

$a_1, a_2 \cdots a_{12}$ be the typical percentage differences between each

month and the average of the adjacent months. The numerical values for the a's in the illustrative example are given in Table II,

TABLE II

MONTHLY FREIGHT REVENUES

	a_1 etc.	(b)* $2M_1a_1$ etc.	(e)*	(g)*		Per cent difference M_1 etc. from R	Seasonal factors
						4.9	95.1
1	.10	—	—	.4	.4	5.2	94.8
2	− .75	− 1.7	−1.7	−1.3	− .9	3.9	95.1
3	2.05	3.9	2.2	2.6	1.7	6.5	93.5
4	−2.30	− 4.6	−2.4	−2.0	− .3	4.6	95.4
5	.75	1.2	−1.2	− .8	− 1.1	3.7	96.3
6	−2.40	− 4.9	−6.1	−5.8	− 6.9	− 2.1	102.1
7	3.45	6.8	.7	1.1	− 5.8	− .9	100.9
8	−5.13	−10.6	−9.9	−9.5	−15.3	−10.5	110.5
9	2.90	6.0	−3.9	−3.5	−18.8	−14.0	114.0
10	5.30	11.7	7.8	8.2	−10.6	− 5.8	105.8
11	1.15	2.1	9.9	10.3	− .3	4.5	95.5
12	−5.10	− 9.9	—	.3	—	4.9	
			−4.6 (f)*		−57.9 (i)*		
			− .3 10/12		− 4.8 3/12		

* Letters refer to equations used in text.

$M_1, M_2 \cdots M_{12}$ be arbitrary values for freight revenues for January, February, · · · December, satisfying the conditions of the a's,

R equal average monthly freight revenues for the whole period.

Then:

The problem is to determine the M's numerically in relation to R as this will give the desired seasonal factors.

The M's are determined numerically from the twelve simultaneous equations of the form:

$$\text{(a)} \qquad \frac{M_1 - \dfrac{M_{12} + M_2}{2}}{M_1} = a_1 \text{ etc.}$$

or, restating set (a)

$$\text{(b)} \qquad \begin{aligned} (M_1 - M_{12}) - (M_2 - M_1) &= 2M_1a_1 \\ (M_2 - M_1) - (M_3 - M_2) &= 2M_1a_2 \text{ etc.} \end{aligned}$$

To solve these equations, arbitrary values must be assigned to the M's used as factors on the right side. The values for the M's differ by only a relatively small percentage from the average freight revenues for the period; therefore the error is not large if R is substituted for

these M's. Solving the equations after this substitution gives an approximation to the various M's in terms of R. The equations are then reworked substituting the M's so determined on the right side. This gives values for the M's which can be used to determine the seasonal factors.

It must be emphasized that this method of solving the equations introduces only a negligible error in the final figure. Seasonal factors derived from the first approximation (using R for all M's on the right side) show a maximum difference of 1.6 points from those derived from the second approximation but factors obtained from a third approximation showed a maximum deviation of only .2 percentage points. The second approximation is accurate enough for all practical purposes and is used in this study.

Substituting numerical values derived as explained above in the right side of set (b) the equations become:

(c) $(M_1-M_{12})-(M_2-M_1)=.00\ R$ etc.
Letting $d_1=M_1-M_{12}$; $d_2=M_2-M_1$; $\cdots$ $d_{12}=M_{12}-M_{11}$
set (c) becomes:

(d) $d_1-d_2=.00\ R$; $d_2-d_3=-.017\ R$; $\cdots$ $d_{12}-d_1=-.099\ R$.
By cumulative addition, this set of equations becomes

(e) $d_1-d_2=.00\ R$; $d_1-d_3=-.017\ R$; $\cdots$ $d_1-d_{12}=+.099\ R$
$d_1-d_1=0$.
Adding set (e):

(f) $12d_1-(d_1+d_2\cdots d_{12})=-.046\ R$
but $d_1+d_2\cdots d_{12}=(M_1-M_{12})+(M_2-M_1)\cdots(M_{12}-M_{11})=0$
so: $d_1=-.003\ 10/12\ R$.
Substituting for d_1 in set (e):

(g) $-d_2=.004\ R$; $-d_3=-.013\ R$ etc.
but $-d_2=M_1-M_2$ etc.

(h) so $M_1-M_2=.004\ R$ etc.
which is in form corresponding to set (d) and can be reduced to:

(i) $12M_1-(M_1+M_2\cdots M_{12})=-.579\ R$
but $(M_1+M_2\cdots M_{12})=12\ R$

(j) so $M_1=R-.048\ 3/12\ R$.
Numerical values of the other M's may be readily obtained.

Table II shows the full numerical derivation of each set of equations. Seasonal factors which minimize the differences of the various M's from R were taken as final and were used to smooth the series.

MOVING SEASONAL INDEXES

By Dudley J. Cowden
University of North Carolina

A WIDELY USED METHOD of computing a seasonal index, whether a stable index or a moving seasonal, is based upon percentages of a twelve month moving average. The moving average is intended as a rough estimate of combined trend and cyclical movements. The moving average, however, tends to smooth out cyclical peaks and troughs. As a consequence percentages of moving average have a downward bias in years containing a sharp trough of depression and an upward bias in years containing a sharp peak of prosperity. In computing a stable seasonal the inaccuracy occasioned by this lack of homogeneity in the data is partly overcome by using a modified mean or median for the monthly arrays.

A moving seasonal is usually obtained by fitting a free hand trend to the percentages for each month separately, and adjusting the trend values so that the total for each year will be 1200, at the same time retaining the smoothness and good fit of each trend. The method is both laborious and subjective, and suffers even more than does a stable seasonal from the lack of homogeneity of the data.

If the percentages of moving average are adjusted (by multiplying by a constant) so that they will total 1200 for each year two objects will be accomplished: (1) the bias due to inflexibility of the moving average will be largely overcome; (2) if polynomial trends are fitted by the method of least squares to each month, the trend values will total 1200 for each year, thus avoiding the adjustment process referred to above. This second result will hold only if the same degree of equation is used for each month.

The reason the trend values total 1200 is that, if an equation of the type $Y_c = a + bX + cX^2 \cdots$ is used, $\Sigma a = 1200$, $\Sigma b = 0$, $\Sigma c = 0$, etc. If one merely wishes to have his seasonal index total 1200 each year, he need not adjust the percentages of moving average to total 1200 before constructing his index. All he need do is to adjust the 12 values of a so that they will total 1200, and the twelve values of each other constant so that they will total zero. This, however, does not overcome the difficulty occasioned by the defect inherent in the moving average, and sometimes the adjustment of the constants will even change the sign of certain constants.

If a time series is deseasonalized by dividing by a seasonal index constructed as described above, the total of the deseasonalized data will

not, in general, be the same as the total of the original data. It would be possible to overcome this difficulty by constructing the seasonal index from differences between the original data and the 12 month moving average, and to deseasonalize by subtracting the seasonal index from the original data. This procedure would be logical if one thinks of the original data as being the sum of the component elements: Trend (T); Cycle (C); Seasonal (S); Irregular (I). According to this concept seasonal movements vary in absolute magnitude according to the trend values and the position of the cycle. A stable seasonal is virtually precluded, however, and moving seasonals are described by unnecessarily complex trends. Consistent with the above concept, but more easy to handle statistically, is to think of the original data as consisting of $T \times C' \times S' \times I'$, the primes indicating that the movements in question are percentages of the combined preceding elements. Also, log original data $=$ log T + log C' + log S' + log I'. If now a logarithmic seasonal index totaling zero is based on the differences between the logarithms of the original data and the moving average of those logarithms, and if seasonal adjustment is made by subtracting the logarithms of the seasonal index from the logarithms of the original data, then the total of the deseasonalized logarithms will be the same as the total of the original logarithms, and the geometric mean of the deseasonalized data will be the same as the geometric mean of the original data. (Also, the geometric mean of the seasonal index will be unity, and the geometric mean of the data deseasonalized by dividing by the seasonal index would be the same as the geometric mean of the original data.) In constructing a logarithmic seasonal index, the difference between the logarithms of the original data and the moving average of those logarithms should be adjusted to total zero each year before computing the seasonal index. If this is done, the total of each of the different trend constants will also total zero.

For the method of computing a moving seasonal index described here, use of orthogonal polynomials will greatly facilitate the computation of the constants, and also the use (with a grain of salt) of covariance analysis for testing the degree of equation to use.

ERRORS IN CARD PUNCHING

By W. Edwards Deming, Benjamin J. Tepping, and Leon Geoffrey
Bureau of the Census

The purpose of this study is to obtain information on errors in card punching, the ultimate object being (i) to provide a basis for determining how such errors, if uncorrected, would affect the Census tabulations, and (ii) to provide a basis for administering sample verification. The introduction of the sample verification[1] of coding and card punching in the Census, and its adoption and extension in large corporations, renders such studies needful. Naturally, this paper must deal with the processing of data in the Census, but many of our conclusions are applicable in other lines of work.

Data collected during a Census enumeration go through many stages of processing before they are summarized in the form of tables for publication. A major stage is the transfer to punch cards of the edited and coded information that appears on the original Census schedules, and on transcription sheets made therefrom. Even the most reliable punch operator occasionally punches a hole in the wrong position on a card, or skips a column, or puts a second hole in a column where there should be only one. It is with these errors that this paper deals.

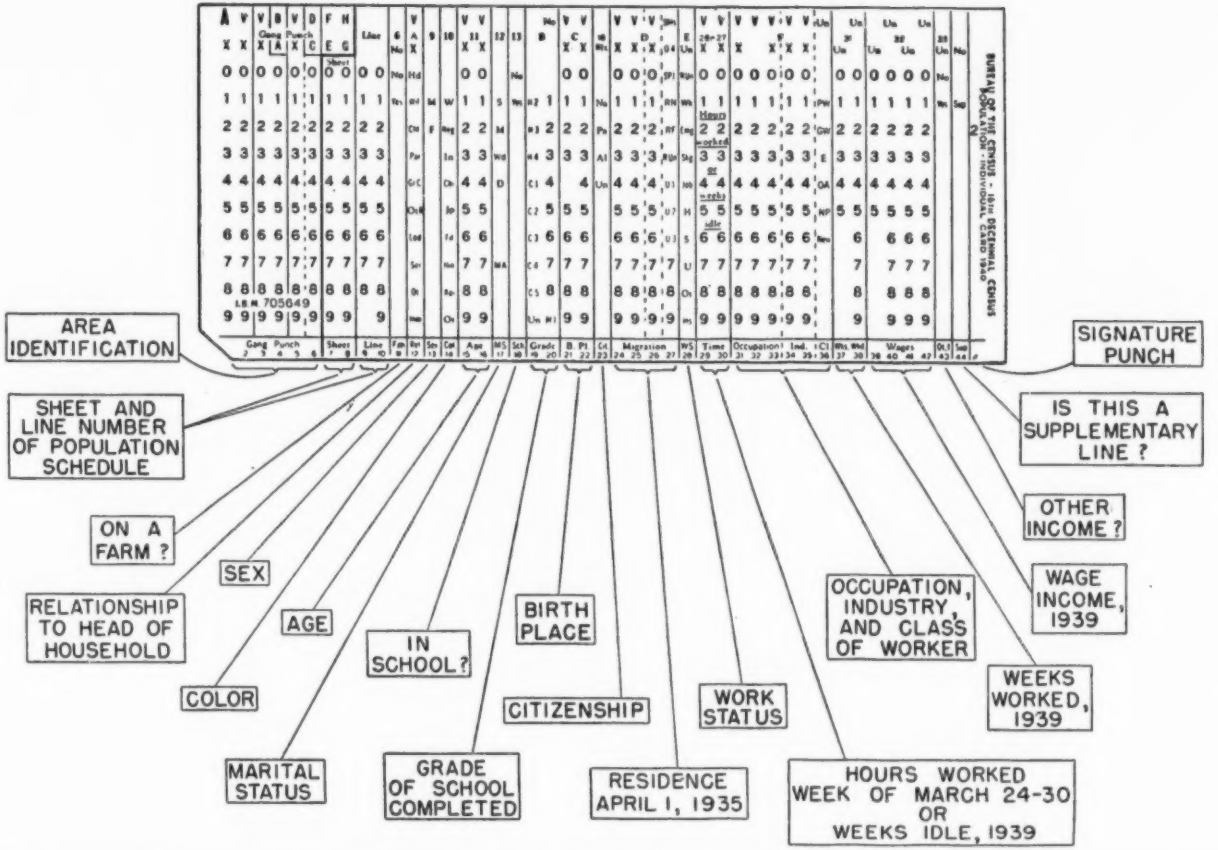
After punching, the cards are sent through the verifying operation. The operator of a verifying machine is expected to discover any error in a card that goes through her machine, whereupon the card is sent back to the original operator who is obliged to punch a new one. Depending on her skill, a punch operator is thus obliged to replace from .5 per cent to 5 per cent of the cards that she punches. The verification and repunching is therefore an item of considerable expense, and any information that will help diminish it should pay high dividends.

The 25,000 study cards. The cards studied were wrongly punched A cards of the 1940 Census of Population. An A card (Chart I) was punched for each person enumerated in the 1940 Census. Five samples, each consisting of 5,000 wrongly punched A cards, were studied. Tests showed that the study cards were sufficiently representative of all wrongly punched A cards to render valid the conclusions drawn in this paper.

[1] See W. Edwards Deming and Leon Geoffrey, "On Sample Inspection in the Processing of Census Returns," this Journal, Vol. 36, September 1941, pp. 351–360. The procedure for sample verification of punch cards, and the extent to which it was used to supplant 100 per cent verification, is described there. Sample verification was introduced only for certain steps in the processing of the Census data, and even then, only for clerks whose work was known individually to be uniformly so good that complete verification and correction of errors would constitute needless and wasteful refinement of their work.

CHART I

THE POPULATION A CARD



THE EFFECT OF WRONG PUNCHES WITHIN SPECIFIC FIELDS OF THE POPULATION A CARD

The effect of wrong punches on the tabulated results. The purpose of this study is to learn how sample verification of card punching affects the tabulations. In sample verification,[1] punchers having a uniform error rate below 1 wrong card per 100 cards punched were "qualified"—i.e., only 5 per cent of their work was put through the verifying machine. The wrong cards in the other 95 per cent went through to tabulation uncorrected. The fact is that even in 100 per cent verification, some errors do get through. The only way to get good cards to tabulate is to have them punched correctly in the first place. Sample verification actually raises the quality of punching by putting a premium on good punching. In the punching of the Population A card (Chart I), and the Housing E card (not shown), at any one time not more than a third of the punchers were "qualified"; and with careful controls used, the qualified punchers contributed no more than their proportionate share of the wrong cards tabulated. Moreover, the study here described shows that when sample verification is used, the net effect of wrong punches, as they occur in practice, is often negligible. There are two reasons for this. First, in many fields of the card (see Chart I), errors tend to compensate for one another. Thus, while some cards were punched male when they should have been punched female, approximately an equal number were punched conversely, and the net error in the tabulations would have been small, even had none of the wrong cards been corrected. Second, it was found that the right-hand columns of a field (the units) are oftener wrongly punched than the earlier ones (tens or hundreds), so that in fields in which a numerical quantity is punched, the errors are mostly of small magnitude. The same pattern fortunately appears in punching fields of two and three digit codes; the right-hand columns are punched wrong oftener than the left-hand ones. The reasons behind this phenomenon are undoubtedly psychological.

Wrong punches in the age field. For cards sample verified, the maximum error rate is 1 wrong card in 100 cards punched. For the cards verified 100 per cent, the proportion of wrong cards sent to tabulation is less. But even if the average error rate is 1 wrong card in 100, wrong punches in the age field would bias the tabulated results very little even if no wrong cards were corrected. Age was incorrectly punched on 3,210 cards among the 25,000 wrongly punched cards. However, 41.8 per cent of these 3,210 were still in the correct 5-year age group, and of the remaining 58.2 per cent, 2 out of 5 were in the adjoining 5-year age groups. On this basis we may make a prediction concerning the maxi-

mum error in the final tabulations arising from sample verification. Of the 25,000 study cards, 1,868 were punched in the wrong 5-year age group. At the rate of 1 wrong card in 100, the 25,000 study cards are all the wrong cards there are for 2,500,000 people, hence only 1,868/2,500,000 or 0.075 per cent of all ages would be classified in the wrong 5-year age group.

This conclusion holds for any cross-tabulation into which age enters. If, however, one is interested only in the distribution of a population by age, the distortion introduced by these errors is even smaller. Thus 428 of the 25,000 cards should have been punched "0 to 4 years," while actually 487 were so punched, a difference of 13.8 per cent. The average amount by which the frequency in any 5-year age group differed from the correct frequency, for the 3,210 cards that were wrongly punched in the age field, was 14.3 per cent. The error introduced in a 5-year age group of the minimum of 2,500,000 cards from which the 25,000 cards are to be considered drawn will thus be less than 0.02 per cent.

Wrong punches in other fields of the card. The pattern of wrong punches in the wage income field is similar to that in the age field. Other fields studied were those for farm residence, color, citizenship, and marital status. The picture is essentially the same for all these fields. Most of the cards were correctly punched, but among the few cards that were wrongly punched, there was a tendency for the errors to compensate for one another. Thus, 24,148 of the 25,000 wrongly punched cards were correctly punched with regard to color; for the 852 cards that were wrongly punched in that field, the distribution by color was not greatly different from what it would have been without the errors.

FACTORS ASSOCIATED WITH THE OCCURRENCE OF WRONG PUNCHES

Clustering of the wrong punches on cards having more than one wrong punch. A "wrong card" has one or more wrong punches. The "error rate" of a puncher is defined as the number of wrong cards per hundred cards punched. Although this method of counting errors has proved administratively convenient, it reveals nothing about the number of wrongly punched columns. It is therefore of interest to examine cards that have wrong punches, in order to determine how many of them have only one wrong punch, how many have two wrong punches, etc.

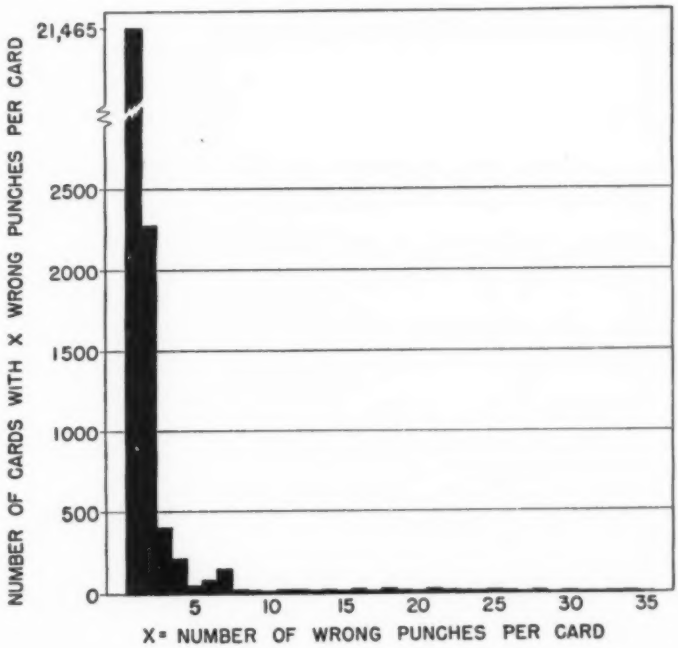
There are 45 columns on a Population A card (Chart I). The machine operated by the puncher is set so that the punching begins with col. 7 and ends with col. 45. The first six columns, which identify the cards by

geographic area, are later punched mechanically. On each card, therefore, there are 39 possibilities for making an error.

Chart II shows the distribution of the 25,000 study cards according to the number of wrong punches per card. Of the wrong cards 86 per cent have only one error, 9 per cent have exactly two errors, and 5 per cent have three or more errors. The average number of wrong punches per wrong card for the 25,000 cards is 1.47. The five subsamples, consisting of 5,000 cards each, exhibited stability about the same value, the averages being 1.51, 1.39, 1.47, 1.54, and 1.42.

CHART II

THE DISTRIBUTION OF THE 25,000 STUDY CARDS ACCORDING TO THE NUMBER OF WRONG PUNCHES PER CARD



Of the 25,000 wrongly punched cards, 3,535 cards contained more than one wrong punch each. An investigation of these "multiple-error" cards showed that often the wrong punches on a card were related. For example, the operator (possibly through fault of the machine) may skip col. 21, whereupon she punches the code for col. 22 into col. 21, the code for col. 23 into col. 22, etc., with the result that all the entries to the

right of col. 20 are wrongly punched, all as the result of the one wrong punch in col. 21.

Multiple errors may be classified as follows:

i. Failure to punch a code belonging in a given column of the card. As a result, each succeeding punch is entered one column to the left of its correct position. This does not mean, however, that every succeeding column is wrongly punched, since the correct punch often appears by coincidence or as a result of a compensating wrong punch.

ii. Failure of the punching machine to move, causing two punches in a given column. This has the same effect as i.

iii. The same code punched in two successive columns. As a result of this, each succeeding punch is entered one column to the right of its correct position.

iv. Interchange of punches. These are instances where the operator has, for example, punched 32 instead of 23, or 545 instead of 454.

v. Wrong punches not related to each other.

Each of the first four types of error results in a card that contains several related wrong punches.

Most errors of types i, ii, and iii are easily identified. On a large proportion of the cards containing errors of types i and ii, the punching ends with col. 44 instead of with col. 45. Such cards therefore have no "signature" punch, which the punching instructions specify must be punched in the "2" position of col. 45 of every card (see Chart I). On the other hand, most cards containing errors of type iii have two punches in col. 45, namely, the punch belonging to col. 44 and the "signature" punch.

The importance of these observations lies in the fact that cards with errors in col. 45 can be separated from the other cards at some stage of the sorting process. Such cards contain many more errors than the average, as is indicated by Table I, which shows the average number of wrong punches per card for various groupings of the 25,000 error cards.

TABLE I

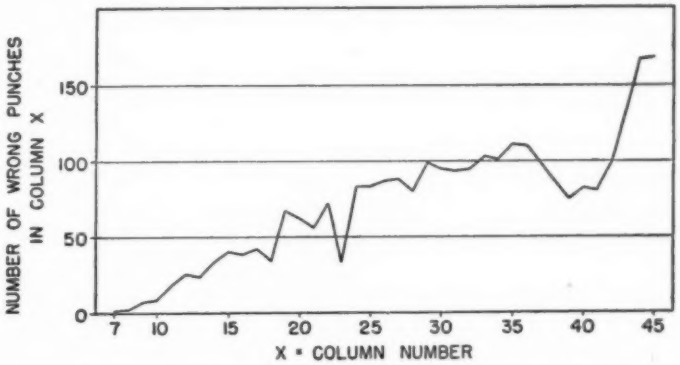
Grouping	Average number of errors per card
25,000 error cards	1.47
3,535 multiple-error cards	4.29
168 cards with no signature punch in col. 45	16.48
24,832 cards with a signature punch in col. 45	1.36

The signature punch in col. 45 was omitted on 168 of the 25,000 study cards. The wrong punches on these 168 cards were concentrated toward the end of the card, as indicated by Chart III, which shows the distribu-

tion of wrong punches according to the columns in which they occur. If the 168 cards are removed from consideration, the distribution (not shown) of wrong punches on the remaining 24,832 cards, according to the column in which the error occurred, obviously changes. The average number of wrong punches per card is reduced to 1.36.

CHART III

DISTRIBUTION BY COLUMN OF 2,769 WRONG PUNCHES ON 168 CARDS HAVING NO "SIGNATURE" PUNCH IN THE LAST COLUMN



Note.—When the signature punch is missing, usually the puncher has somewhere skipped an entry that was to be punched, or (through fault of the machine) has punched two successive entries into one column of the card.

Cards with errors of type iv constitute about 24 per cent of the 3,535 multiple-error cards among the 25,000 studied. None of these involved more than four columns, and most of them involved only two columns. The proportion of cards with this type of error is very nearly the same in each of the five subsamples of 5,000 cards.

Among the 3,535 multiple-error cards there were 389 of type v. On only 6 cards of the 389, three wrong punches were apparently unrelated, and no card contained more than three unrelated wrong punches. During the period that these cards were punched, an average of 3 cards was wrongly punched out of every 200, or 1.5 per cent. On the basis of simple probability, we might therefore expect about 1.5 per cent of the 25,000 study cards to have 2 or more unrelated wrong punches, and about .0225 per cent of them to have 3 or more unrelated wrong punches. Table II shows the agreement between this theory and actual observation.

The distribution of wrongly punched cards according to schedule line number. Chart IV shows the distribution of the wrong cards according to line number of the population schedule. The population schedule

TABLE II

Cards	Theoretical number	Observed number	Difference	Standard error of the difference
Cards with 2 or more unrelated wrong punches	375.0	389	14.0	19.2
Cards with 3 or more unrelated wrong punches	5.6	6	0.4	2.4
Cards with 4 or more unrelated wrong punches	0.1	0	0.1	0.3

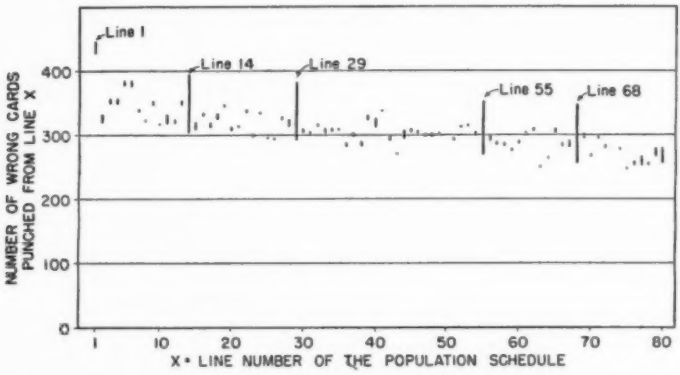
contained 80 lines, i.e., spaces for 80 names, 40 on one side and 40 on the other. Two facts are apparent from Chart IV:

i. There is a gradual decrease in the number of wrong cards for the successive lines of the population schedule.

ii. The number of wrong cards for certain lines is considerably greater than for the others.

CHART IV

THE DISTRIBUTION OF THE 25,000 STUDY CARDS ACCORDING TO LINE NUMBER OF THE POPULATION SCHEDULE



Note.—The ordinate of the upper end of each vertical bar represents the number of wrong cards punched for line x of the Population Schedule. The ordinate of the lower end of each vertical bar represents the number of wrong cards punched for line x, after the 806 cards wrongly punched in col. 44 were excluded. The length of each vertical bar represents the number of cards punched for line x, and having wrong punches in col. 44. The slight downward trend arises from the fact that line x of the schedules was filled more frequently than line $x+1$.

The gradual decrease is to be expected since, on the last schedule that is filled by each enumerator, the early lines are oftener filled than the later ones. Other studies have shown this, and it need not be of concern here. The prominence of the five lines 1, 14, 29, 55, and 68, is, however, more significant. They are sufficiently outstanding to indicate the presence of assignable causes of error. Lines 14, 29, 55, and 68 will be discussed as a group, followed by remarks on line 1.

Lines 14, 29, 55, and 68 are the lines that were designated for supplementary questions on 80 per cent of the schedules.[2] Col. 44 of the punch card is the only column for which the punch is dependent upon the designation of a supplementary line as supplementary. The puncher was instructed to punch "1" in this column for supplementary lines, and to punch "X" for all other lines. Since the lines for which "1" is to be punched constitute only 5 per cent of all lines, it seems likely, on the basis of related coding and punching experience, that the puncher would occasionally by habit punch the usual "X" on a card that required "1."

As a test, then, the 25,000 wrong cards were divided into two groups:

i. Cards that were punched for lines 14, 29, 55, and 68;
ii. Cards that were punched for all other lines.

In the first group, one out of every four cards had an error in col. 44, whereas in the second group only one out of fifty had an error in this column. Furthermore, the number of cards with errors in col. 44 for any one of the four lines was at least 4 times the number of cards with errors in col. 44 for any one of the remaining 76 lines of the schedule. Chart IV shows that when cards with errors in col. 44 are excluded, lines 14, 29, 55, and 68 are no longer prominent. We conclude, therefore, that the excess of wrong cards for lines 14, 29, 55, and 68 is due to errors in col. 44.

The errors on cards punched from line 1 of the population schedule do not appear to differ in kind from errors on the other cards. Unlike the group of wrong cards for lines 14, 29, 55, and 68, no particular column seems to account for the errors on the cards for line 1. Now line 1 is the first one that is punched after a break in the puncher's rhythm (removing a completed schedule, and inserting a new one in the holder), and if this is the correct explanation, one would expect to find more than the usual number of errors also for the following types of cards:

i. Cards for line 41. Before punching line 41, the puncher must turn the schedule over in the holder.
ii. Cards for line 1 of the first schedule in a folio. The puncher stops to hand in completed work before commencing a new folio.

With regard to line 41, an excess does occur (see Chart IV), although it is not as much as for line 1. This seems reasonable: the interval is shorter for line 41 than for line 1. With regard to the second point, the interval (time out) preceding line 1 is longer for the first schedule in a

[2] Frederick F. Stephan, W. Edwards Deming, and Morris H. Hansen, "The Sampling Procedure of the 1940 Population Census" this JOURNAL, Vol. 35, December 1940, pp. 615–630.

folio than for any other schedule in the folio. This longer interval was indicated as the cause of the prominence at line 1 when it was found that there were more than twice as many wrong cards from line 1 of schedule 1 as from line 1 of schedule 2 (the schedule that contributed the second largest number of wrong cards).

The frequency of a code in relation to wrong punching. Early in the course of this study, there appeared a principle that seems basic. For some areas, and for some fields of the Population A card, there may occur what we call a "usual" punch. This is a punch which occurs much more frequently than any other punch in the column or field. For example, the "usual" punch in Field 16 (citizenship) is "native," and the other possible punches in this field are relatively rare. When a punch

TABLE III

Errors in Field 10 (Color)

	Color to be punched		
	White	Nonwhite	Total
Cards correctly punched in Field 10	21,029 (98.0%)	3,119 (88.1%)	24,148
Cards incorrectly punched in Field 10	431 (2.0%)	421 (11.9%)	852
Total	21,460	3,540	25,000

Errors in Field 12 (Marital Status)

	Marital Status to be punched		
	Single or married	Any other status	Total
Cards correctly punched in Field 12	20,565 (94.5%)	2,924 (87.5%)	23,489
Cards incorrectly punched in Field 12	1,093 (5.5%)	418 (12.5%)	1,511
Total	21,658	3,342	25,000

Errors in Field 16 (Citizenship)

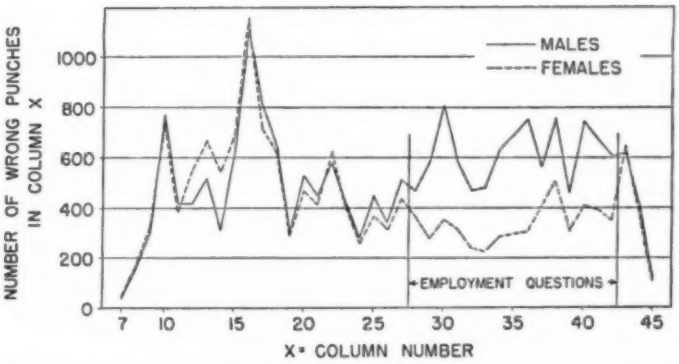
	Citizenship to be punched		
	Native	Any other citizenship	Total
Cards correctly punched in Field 16	22,360 (98.9%)	1,814 (76.2%)	24,174
Cards incorrectly punched in Field 16	260 (1.1%)	566 (23.8%)	826
Total	22,620	2,380	25,000

operator encounters one of the rare entries, she may from habit simply punch it as usual. Thus we expect to find a larger proportion of wrong punches among those cards on which "unusual" entries should be punched. This expectation is borne out by our study of the 25,000 wrongly punched cards.

Table III demonstrates the operation of the principle. In Field 10 (color), "white" may be considered the usual punch. Of the 21,460 cards punched for white persons, only 2.0 per cent were incorrectly punched in the color column; of the 2,540 cards punched for nonwhite persons,

CHART V

THE DISTRIBUTION OF THE WRONG PUNCHES ON THE 14,131 STUDY CARDS PUNCHED FOR MALES AND ON THE 10,869 STUDY CARDS PUNCHED FOR FEMALES



11.9 per cent were incorrectly punched in the color column. Thus it appears to be more than 5 times as easy to make an error in the color column for a nonwhite person as for a white person.

A similar situation holds in the case of Field 16 (citizenship). Among the 25,000 study cards, only 1.1 per cent of those for native Americans showed the wrong citizenship, but 23.8 per cent (or, relatively 20 times as many) of those for persons not native Americans showed the wrong citizenship.

The results of the principle just described are sometimes not so obvious. It will be noted that 14,131 of the 25,000 cards were punched for males. This large proportion is easily accounted for, when it is observed that cards punched for females will often have a "usual" entry in the employment fields (cols. 28–43, inclusive). These "usual" punches are "unknown" and "zero." Males will not so often have these entries. Investigation showed that the larger number of cards for males was a con-

sequence of wrong punches in the employment columns (see Chart V).

In the same way that the action of the principle described above results in an under-representation of females among the wrong cards, it also results in an under-representation of children under 14 years, since they also will have "usual" entries in the employment fields.

In the preceding discussion we have shown that an "unusual" entry is oftener punched incorrectly than a "usual" punch. It was further determined that when an "unusual" entry is punched incorrectly, it is much more likely to be punched as the "usual" punch (if any) than as any other punch.

The actual labor of drawing the sample (the 25,000 study cards), examining them, compiling and studying the figures, was carried out by various staff members, most of it by or under the immediate supervision of Mr. Samuel W. Greenhouse.

CORRELATIONS BETWEEN FUNCTIONS OF VARIABLES

By G. A. Baker
University of California at Davis

The question of the correlation between functions of variables has been raised by Dr. A. E. Brandt.[1] He considers the simple case of the correlation between x_1 and (x_1+x_2) and erroneously comes to the conclusion that such a correlation can never be zero and that the ordinary tests of significance for correlation coefficients cannot be applied. The purpose of this note is to show that the question of correlating a part with a whole is a very restricted case of a very general situation that is of wide practical and theoretical interest and that in many cases the correlations between functions of variables should be computed and compared with zero or their expected value by means of the usual tests of significance.

Sometimes a deliberate search is made for a function of one or more variables that is uncorrelated with another function of one or more variables. As an example, we shall consider the correlation between x_1 and (x_1+x_2) and show that the correlation can be zero in this case. We shall, for the moment, consider only variables that are normally correlated. Consider the populations

$$Z = \frac{1}{2\pi\sigma_1\sqrt{\sigma_2^2-\sigma_1^2}} \exp\left(-\frac{\sigma_2^2}{2(\sigma_2^2-\sigma_1^2)}\left(\frac{x_1^2}{\sigma_1^2}+\frac{x_2^2}{\sigma_2^2}+\frac{2x_1x_2}{\sigma_2^2}\right)\right), \qquad \sigma_2 > \sigma_1. \quad (1)$$

Thus x_1 and x_2 are two normally correlated variables with their correlation coefficient equal to $-\sigma_1/\sigma_2$. The populations (1) are not isolated, trivial ones but consist of an infinite class, i.e., of populations for all values of σ_2, σ_1 below the line $\sigma_2=\sigma_1$ in the first quadrant of the (σ_2, σ_1) plane. Now make the transformation of variables

$$u = x_1 \qquad x_1 = u \qquad dx_1dx_2 = dudv$$
$$v = x_1 + x_2 \qquad x_2 = v - u$$

and (1) becomes

$$Z = \frac{1}{(2\pi\sigma_1\sqrt{\sigma_2^2-\sigma_1^2})} \exp\left(-\tfrac{1}{2}\left(\frac{u^2}{\sigma_1^2}+\frac{v^2}{\sigma_2^2-\sigma_1^2}\right)\right), \qquad \sigma_2 > \sigma_1. \quad (2)$$

[1] "Should We Correlate a Part and the Whole?" Statistical Service Sheet No. 2, U.S.D.A. Office of Information, Division of Publications, Washington, D. C., October 10, 1941.

The variables u and v are independent. Since u is a "part" of v we have an example of zero correlation between a "part" and a "whole" which consists of a whole class of populations. In sampling from such populations as (1) it is proper to consider the "part"-"whole" correlation and test its divergence from zero by means of the usual significance tests.

In general, consider two variables x_1 and x_2 measured from their means, with standard deviations σ_1 and σ_2 and product correlation, r. Then the expected value of x_1 $(x_1+x_2)/\sigma_1\sqrt{\sigma_1^2+\sigma_2^2+2r\sigma_1\sigma_2}$ is $(\sigma_1+\sigma_2 r)/\sqrt{\sigma_1^2\sigma_2^2+2r\sigma_1\sigma_2}$ which is zero if $r=-\sigma_1/\sigma_2$ and $\sigma_2>\sigma_1$. In all such cases it is proper to test divergence of sample "part"-"whole" correlations from zero. Such a situation might arise in case two characters were negatively correlated but one character was more variable than another.

The correlation of x_1 with (x_1+x_2) has been discussed at some length because it is a simple example of "spurious" correlation and illustrates some of the considerations in connection with the use of "spurious" correlations. "Spurious" correlation was first discussed by Karl Pearson.[2] Yule[3] takes a different view of "spurious" correlation. He says, "It does not follow of necessity that the correlations between indices or ratios are misleading. If the indices are uncorrelated, there will be a similar "spurious" correlation between the absolute measurements $z_1x_3=x_1$ and $z_2x_3=x_2$, and the answer to the question whether the correlation between indices or that between absolute measures is misleading depends on the further question whether the indices or the absolute measures are the quantities directly determined by the causes under investigation."

Many modern investigators have taken the view of Yule. As an example, W. L. Gaines, C. S. Rhode, and J. G. Cash[4] considered the three variables

$x_1=FCM$ (eight-months partial lactation milk-energy yields in terms of pounds of 4 per cent milk per day)

$x_2=$ age in years

$x_3=$ initial live weight in pounds.

In this case $r_{12}\neq 0$, $r_{13}\neq 0$, $r_{23}\neq 0$. The investigators deliberately sought a function of x_1 and x_3 that would not be correlated with x_2. Such a function turned out to be x_1/x_3. To test whether the variable x_1/x_3 was or

[2] "On a Form of Spurious Correlation which may Arise when Indices Are Used in the Measurement of Organs," Proceedings of the Royal Society, Vol. LX, 1897, p. 489.

[3] "An Introduction to the Theory of Statistics," C. Griffin & Co., Ltd., London, 10th edition, revised (1932), p. 216.

[4] "Age, Live Weight and Milk-Energy Yield—A Correction," *Journal of Dairy Science*, Vol. 25, No. 1, 1942, pp. 15–18.

was not correlated with x_2 the deviation of the resulting correlation coefficient from zero was considered.

In general, instead of discussing the correlation between x_1 and (x_1+x_2) we might talk about the correlation between x_1 and some other expression in x_1 and x_2, say $f(x_1, x_2)$. To generalize still further, we might consider the correlation between $g(x_1, x_2, \cdots)$ and $f(x_1, x_2, \cdots)$ where $x_1, x_2, \cdots$ are some measured variables which may or may not be correlated. In the discussion of Yule and Pearson $g(x_1, x_2, \cdots)=x_1/x_3$ and $f(x_1, x_2, \cdots)=x_2/x_3$. In the dairy example $g(x_1, x_2, \cdots)=x_2$ and $f(x_1, x_2, \cdots)=x_1/x_3$. In some cases we may be interested in the correlations between the x_i's and then the correlations between g's and f's might be termed spurious. In other cases we may be interested in the g's and f's or in the particular g's and f's that are not correlated and then the correlations between the x_i's would be termed spurious. To test whether we actually have g's and f's that are uncorrelated we test the significance of the deviation of the observed correlation from zero.

It has been shown that the point of view and purposes of the investigator determine whether or not "part"-"whole" correlations should be computed and the significance of their deviations from zero be tested. To condemn utterly the use of "part"-"whole" correlation is an error that should not be allowed to creep into authoritative statistical writings to add to the confusion of the statistical practitioners. Even though the expected value of a "part"-"whole" correlation is not zero it may still be of value to compute such correlations if the investigator has clearly in mind what he is doing. For instance, in a recent investigation of asparagus at the Davis, California, Experiment Station the "part" yields for 1, 2, $\cdots$, 9 years were correlated with the "whole" 10-year yields for different classes of plants. The expected value of the correlations, if no correlation existed between the yields for different years and the standard deviations were the same for different years, was easily seen to be $\sqrt{x}/\sqrt{10}$, $x=1, 2, \cdots, 9$. The observed values of the "part"-"whole" correlations for the different classes of plants behaved differently with respect to these expected values. This difference in behavior was of great practical and theoretical interest.

A COMMENT ON DEMING'S CLASSIFICATION OF PROBLEMS OF INFERENCE

By Leonard A. Salter, Jr.
University of Wisconsin

Mr. Deming's article in the June issue of this Journal contains numerous observations of great importance to research workers, and especially to social scientists. No major exceptions can be taken to the thesis which is set forth. It might be suggested, however, that the implications of the article are much greater even than the paper would lead one to believe. Furthermore, for social scientists at least, the central import of Mr. Deming's statement is that it exposes a basic error in the type of research procedures which have been emphasized during the past twenty years, and it calls for a full reappraisal of customary techniques.

The Type A "problems" to which Mr. Deming refers are actually but problems of *counting*. They are not problems the solution of which is necessary in order to know what to do in order to put action under purposive control. They are questions of counting those elements in a static situation that have previously been determined as strategic, in order to apply an already-decided-upon remedy. The crucial points in the illustrations of Type A social "problems" (the allocation of funds by states or conscription by age groups) are these: that the existential social problem has somehow been defined beforehand, and the basic solution of it has already been decided upon. All that remains for the statistician to do is to count out the elements involved for administrative purposes.

As Mr. Deming says, it is only Type B analyses that can show how a given condition "got that way, or what it ought to be or might have been." Some people would, with excellent reason, insist upon adding that it is just these questions which constitute scientific inquiry. One can only rejoice over Mr. Deming's basic introductory remark that scientific data are collected, not for museums, but to provide recommendations for action. But this is also saying that a science moves forward on Type B achievements. The point about measuring Mr. Deming's table is not merely that in stating its measurement one is predicting that its length will be six feet: the important thing is that a *scientist* measures the table because he is trying to do something with the table. And this means that as a scientist he is really interested in some Type B problem and not merely in the Type A process of measuring the table.

The investigational procedures which Dr. Shewhart classifies as

"classical statistical theory" are the procedures that are directly applicable to the kind of "problems" which Mr. Deming classifies as Type A. Now then, when the very limited nature of Type A "problems" is fully grasped, social scientists are found to be in a pretty fix. The resolution of a confused situation into a defined problem and the resolution of a defined problem into a solution that can be used in purposive action—these are inherently what Deming labels as Type B problems. Type A questions become a matter of interest primarily when the problem has been defined and the solution already accepted, and this point can be achieved only by Type B investigations.

Mr. Deming is of course quite correct in asserting that Type A procedures can often help to establish partial evidence for a Type B study, but this hardly needs to be underscored today. What rather needs to be emphasized more is that procedures premised on a Type A concept of problems cannot alone give grounded solutions to Type B problems. On this issue, the physical and social scientists stand far apart in practice. The physical scientist can usually operate with his hands and eyes to observe the all-important sequences in his data. He may write up his results so that all his work looks like Type A. But in the laboratory, he operates in the manner of Type B. As John Dewey has said, we should pay more attention to what the physical scientist does than to what he says he does.[1] This is the essence of Mr. Deming's illustration about soil treatment experiments.

In the field of social science, the research worker cannot easily observe operating processes at first hand. He has to find instances where various social actions have gone through a certain cycle and he has to find the on-going processes of the present. Because these processes may be clearly revealed and exposed at different times and places, because many of them take a long while to unwind, and because many of them are not directly visible to the eye anyhow, the social scientist has to lean heavily on written symbols representing the action in which he is interested. But just because the social scientist makes surveys is no reason for presuming that these are Type A analyses. If they are only Type A analyses, the public needs to be warned of the strict limits on the usefulness of the results. The social scientist uses surveys as often as he does because he has to. Nevertheless, he still has to meet the requirements of Type B analysis by handling his data, however obtained, in a Type B manner. Thus the question of developing Type B procedures for social studies assumes dominant significance if there is to be any science in social science.

[1] On most of these questions of scientific inquiry, John Dewey's *Logic, the Science of Inquiry* will be found helpful if not essential.

In my opinion, Mr. Deming's phrase, "repeated patterns in subseries" reflects the kernel of the procedural methods which need to be developed in social science. The social scientist cannot merely refer to "the underlying forces (in a Type B problem) however derived" even though the statistician may be able to. Nor will it be enough to recommend quick and frequent small samples as long as Type A procedures themselves can only ground solutions of Type A problems, because scientific inquiry has a Type B heart.

It is to be hoped that Mr. Deming's paper will help to force more attention to the basic logical problems of social inquiry and to the development of suggestive aids for the handling of social data in such a way as to preserve the all-important patterns of attributes within the observed cases.[2] If social science is to be able to achieve what Dr. Shewhart has made so much progress in doing for engineering problems, there is plenty of room for those who may care to work on these fundamental problems.

[2] A recent effort in this direction by the writer is: "Cross-Sectional and Case-Grouping Procedures in Research Analysis," *Journal of Farm Economics*, Vol. XXIV, No. 3, November 1942.

CARROLL W. DOTEN, 1871–1942

The death of Professor Carroll W. Doten removed from the roll of the Massachusetts Institute of Technology the name of a distinguished and beloved member of that body. During his long career as a teacher in his chosen fields of economics and statistics, he not only established himself in the hearts and minds of the thousands of students who passed through his classes, but also attained a place of honor among his colleagues on the staff. The wide range of interests which claimed Professor Doten's thought and effort led him into many activities, where he served well his community, his state and his country. For Cambridge, he gave service as president of the Associated Charities, president of the Park Commission and president of the Cambridge Club, among many other similar efforts. For Massachusetts, he rendered service as a member of the bureau of immigration, and in numerous other agencies. For his country, he participated in the work of the United States Shipping Board Emergency Fleet Corporation in 1918 and in the work of the United States Census Bureau. In his professional capacity he contributed to the united efforts of both economists and statisticians. He did research work for the Boston School for Social Workers, was Chief Investigator for the Massachusetts Commission on Compensation for Industrial Accidents, was consulting economist for the National Retail Dry Goods Association and was president of the American Statistical Association.

A native of Vermont, Professor Doten retained throughout his entire life a deep devotion to the state of his birth and to the University of Vermont, from which he received the Bachelor's degree in 1895. During the closing years of his life, following his retirement from Technology, he returned to his native state, to enjoy the evening of his days among the associations of his youth. There in his seventy-first year, he passed away, leaving a great circle of friends, to whom his genial smile and encouraging word had been a joy and inspiration for many years.

FLOYD E. ARMSTRONG

BOOK REVIEWS

GLENN E. MCLAUGHLIN
Review Editor

The Long-run Behavior of Costs in a Chain of Shoe Stores: A Statistical Analysis, by Joel Dean and R. Warren James. Studies in Business Administration, Vol. XII, No. 3. Chicago: The University of Chicago Press. 1942. x, 54 pp. $1.00.

This book attempts to determine statistically the effect of changes in the size of retail shoe stores on their operating costs. This knowledge is sought partly in understanding the theory of long-run cost behavior of an individual firm. Administratively, it is sought for the help it might afford in the formulation of a rational policy of store size and location.

The study is not a long-term analysis as might be expected from the statement of aims but a study of the operating costs of fifty-five (out of several hundred branches of a retail shoe concern) selling men's, boys' and girls' shoes during 1937 and 1938. The authors dismiss the idea of making an investigation of the effect on costs of long-term changes in the level of output partly because such an analysis is "likely" to cover only minor changes in size and "seldom large-scale decreases (changes?)." This seems tantamount to rejecting something for being what it is.

The authors contend that the relation between size and cost of retailing by a group of shops selling the same product and identical in every way except size can be considered as the equivalent of a series of successful adaptations by a retail shop in size and costs to changes in the volume of demand. Perhaps because all of the shops are in one city, the authors fail to deal with the objection that store location might affect the validity of their position although later they show that the daily and weekly sales loads of business-district and residential-district stores in the same city differ considerably. The two authors argue that they need not exclude the business district shops from their analysis as their inclusion only weakens the force of their findings. Apparently, they do not realize that they are dealing with two distinctly different types of retail shop problems whose costs have little relation to each other. They also fail to deal with the objection that a retail chain may often keep poor-paying shops open because of the good will and indirect advertising value they may have. Moreover, the method of allocating central costs such as administrative, insurance advertising and indirect general expenses, rent, and some window display costs tends to hide the true burden of operating the small marginal shops. Where such costs are allocated in proportion to the sales of each shop to the total, some of the economies of bulk purchases and large-scale operation (largely contributed by the bigger shops) are transferred to all of the shops equally.

The most important objection to the authors' method of conducting the statistical analysis is the fact that they chose "a group of stores selling differ-

ent proportions of various types of shoes" whose sales as a total for all shops was "representative of the sales volume of the stores in the chain as a whole." To the reviewer, this would seem to have destroyed their basic assumption that the products of the stores were homogeneous: they have deliberately sought heterogeneity.

The physical volume of sales is chosen as the unit of measurement of shop sales although this measure excludes the sales of rubbers, hosiery, and sundries. In choosing this scale it is recognized that it gives equal weight per pair of shoes sold to the time spent in selling men's shoes and the cheaper shoes for juveniles, although the latter class of sales consumes a larger amount of time per sale. The reviewer feels that if this unit of measurement is adopted, it should exclude boys' and girls' shoe sales and cover men's shoe sales only. However, the reviewer regards the physical volume of shoe sales (in the absence of physical volume figures for rubbers, hosiery, and sundries) weighted by the average time spent in selling each pair as a much better index particularly where, as in this case, heterogeneity in the nature of sales between shops is deliberately sought. It would not have been a difficult matter to have made studies of the average amount of time spent in selling each type and style of shoe and to confine the study to those stores whose salesmen spent approximately the same amount of time in making each type of sale.

In measuring cost, two elements of total cost are excluded as tending "to obscure the relations which it is desired to investigate." One of these elements is general administrative expense, the other general indirect expense. The method used in allocating these costs is not stated, but if the volume of sales was the basis then for the sake of consistency, advertising, insurance, rent, and a part of window display charges ought to have been excluded as well.

Correlations are drawn between the physical volume of shoe sales and each element, groups of elements, and total operating costs. The results show that "the elasticity of total costs is increasing over the whole range of output," and is due almost entirely to salesmen's salaries. Some of the groupings of components are open to serious questioning.

In the theoretical section, an attempt is made to explain why the empirical cost function does not bear out the traditional umbrella-shape explanation beloved by theorists explaining the relation between costs and output. The function is shown to consist of an ascending series of discontinuous planes. The explanation of the cost line is first considered as temporary departures from an umbrella or parabolic curve and rejected as completely unrelated to any curve series. The shape is explained as the result of discontinuities in costs arising from increases in output beyond the optimum output at that price. Any attempt to determine the shape of the cost line for decreases in output below each optimum level is expressly disclaimed.

Michael Daly

Washington, D. C.

Home Propaganda, Change No. 2. A Report prepared by Mass-Observation for the Advertising Service Guild. London: The Advertising Service Guild. 1942. vi, 78 pp. One shilling.

This stimulating booklet presents an analysis of the effectiveness of propaganda directed by the British government toward the home front during the first two years of the present World War. The study was prepared for an association of advertising agents by the staff of Mass-Observation, directed by the energetic and resourceful Tom Harrisson. Although some of the basic facts were secured from the numerous part-time investigators who cooperate with the Mass-Observation organization, most of the primary data were obtained by whole-time workers who made door-to-door calls in London, Worcester, and a group of Kentish villages—324 interviews in each area.

According to this investigation, large governmental expenditures, including over half a million dollars a month for poster and newspaper advertising alone, often had negligible results and occasionally even developed antagonism rather than cooperation. As might be expected from observation of similar activities in the United States, British propaganda campaigns are frequently confusing and even contradictory, as when a single issue of a farm journal contains a plea from one governmental agency for farmers to sell their straw for paper making, along with advice from the ministry of agriculture to pulp the straw for feeding cattle.

The investigation also indicates the dangers of propaganda saturation and the relatively greater returns from appeals "at the patent medicine level" than from those expressed in the dignified language of officialdom. In a democracy at war, it is pointed out, effective propaganda becomes a necessity for persuading people to follow their leaders quickly and willingly.

The report does not, however, over-emphasize the significance of propaganda. "The propaganda of events is more potent that the propaganda of words." A healthy scepticism is also exhibited concerning interviews, since they reveal "what is on the public tongue rather than in the public mind."

On the other hand, the study does reveal a lack of awareness of problems of sampling and reliability. For example, no mention is made of the fact that 324 interviews is a more adequate sample for the "group of Kentish villages" than the same number is for London. Similarly, the evidence that publicity campaigns actually decreased knowledge of governmental programs should be accepted with caution in view of the usual willingness of the Englishman to cooperate, even though grumblingly, with officials. Obviously, Mass-Observation would benefit from having a competent statistician on its staff.

Yet in spite of its possible errors of reliability this study does reveal the futility of much government propaganda with its ill-timed and conflicting appeals. As in other publications by Mass-Observation, the data are presented with clarity and vigor. The report should be carefully studied by social scientists interested in the control or measurement of public opinion

and especially by the officials responsible for the present flood of governmental propaganda for the guidance of the American people.

H. C. Brearley

George Peabody College

Parity, Parity, Parity, by John D. Black. Cambridge: The Harvard Committee on Research in the Social Sciences. 1942. xi, 367 pp.

This book treats not only of "parity" between agriculture and non-agriculture, but of price policy in wartime and peacetime as well. The manuscript left the author's hands last May. But the substance remains timely, and will continue so for as long as the parity argument continues to be used in support of legislation aimed to enhance farm prices and farm income. The twenty-four chapters and two forewords are packed with historical facts noted at first hand, replete with ingenious statistical attempts to calculate rational parity standards, full of the shrewd observations on men and events that we have come to expect from Dr. Black. Style and organization—including italics and asterisks—are deliberately journalistic, and to this reviewer more confusing than helpful; but the index helps. Some of the charts are not easy to read on the printed page. The chapters on "Alternative Parity Standards" and "Parity, Please" come closest to summaries of conclusions and warrant the closest reading, although the author describes his conclusions as "relatively the least valuable part of the effort."

Dr. Black is apparently troubled because notions about parity, so useful in gaining support for legislation that he would no doubt call "constructive" in times of peace, are now obstructing the war effort. His concern is justified. If anything is needed in furtherance of war effort now, it is sub-parity for some producers—of wheat, for example—and super-parity for others.

The author apparently dreads inflation more than the reviewer does. The general price freeze of April 28 is spoken of as "right in its objectives." The only reasonable price objective to the reviewer is inflation at a rate as slow as can be engineered while stimulating production for war. If painful postwar deflation must follow, it is nevertheless preferable to defeat. Production for war comes first, postwar deflation or not, parity or not.

Obviously the author does not agree with President Roosevelt in regarding parity for farmers based on 1910–14 as an "excellent objective," now or earlier. He suggests other bases as preferable (chapter xii); and so they are if parity must indeed be lived with. He "holds to the judgment that Agriculture during the whole 72 years has been underpaid relative to both Labor and Capital." But he fails to point out that such a situation is inevitable and desirable in the course of a nation's rise from poverty to riches. There must be an incentive to enlarge the proportion of population engaged in non-farm occupations; and where is the incentive to be found unless real incomes are smaller in farming than out?

Obviously also, Dr. Black believes that "the essential principle of parity is sound." In the reviewer's opinion it is not, even in times of peace. The notion that "parity" between "agriculture" and "non-agriculture" ought to be an appropriate object of national economic policy is in fact the second biggest gold brick that has been sold to the American public in many years. Only protectionism is a bigger one—and it is the parent of parity.

In his concluding remarks on "Parity, Please," the author observes, "It is more than possible that within a year an aroused public will wipe the slate clean of all parities . . . " This is a consummation devoutly to be wished. It seems hardly to be expected, however, unless the public is willing to forsake its current idol of "equal returns for equal effort," returning to the older ideal of remuneration for individuals according to what society will pay for goods and services produced in a competitive system where property rights are respected, or fixing its affections on the current aim of getting the necessary goods produced so as to win the war.

M. K. BENNETT

Food Research Institute
Stanford University, California

The Control of Customer Returns, by Edgar H. Gault and Charles S. Goodman. Ann Arbor: University of Michigan, School of Business, Bureau of Business Research. 1942. 107 pp. $1.50.

Customer returns is an ever present problem in the retail field. Buyers blame salespeople for poor selling; salespeople maintain that customers should make up their minds when purchasing; customers criticize management if merchandise is refused for credit; and management blames buyers, salespeople, and customers when returns become an issue. To pin the tail on the donkey when blindfolded to all the facts is recognized by any retailer as a difficult task.

The Control of Customer Returns is a one-hundred page monograph in which the facts surrounding the return of merchandise are made clear. The study does three things. First, it discusses the nature of the return problem under the familiar headings of procedures in handling returns, causes, and costs. The relation of returns to markdowns and the philosophy of the approval sale add new substance to an old subject.

Second, it indicates the policies for handling returns by calling attention to each store's responsibility in doing a better operating job. These policies, four in number, may or may not be a part of a city-wide plan but should be the basis on which an individual store operates.

Third, it discusses some of the procedures that can be used by stores in controlling customer returns.

Customer returns is not a new subject, but this study has a realistic point of view which should attract more than passing attention. Many retailers and commentators consider customer returns to be a liability. In this report

the liability is acknowledged, but emphasis is placed on "the liberal offering of the return privilege as a powerful promotional device. As a fundamental part of the merchandising operation of a store this policy is expected to increase net profit, while skilled management keeps returns under control through the various techniques which have been described and evaluated." Examples are given to substantiate the argument.

Controlling returns, not attempting to eliminate them because the store shares some of the blame, maintaining that cooperative effort is of questionable value and that each store has its own job to do are important subjects in the report.

Objections to the fact that the study centered in Detroit should not be considered. The work done in other cities has been weighed carefully and the authors' understanding of the subject is too comprehensive to be discounted on this basis.

Here and there a statement is open to question. "The differences observed between the proportion of charge and cash sales returned should not be attributed entirely to the fact of credit purchasing" is controversial. Many purchases are charged which could never be made if they were paid for at time of purchase. Again, "from our own point of view only, the typical customer is not likely to make purchases which are certain to be returned" might be construed as an over-eagerness to prove a point.

None of these criticisms detracts from the real value of the study. Some merchandise men will not agree with the report, but if a retailer who has long considered himself a martyr to customer returns will practice the policies suggested he will have integrated an operation in his business which should have been a part of it long ago.

Bishop Brown

Research Bureau for Retail Training
University of Pittsburgh

An Objective Method of Sampling Wheat Fields to Estimate Production and Quality of Wheat, by Arnold J. King, Dale E. McCarty, and Miles McPeek. U. S. Department of Agriculture, Washington, D. C. Technical Bulletin No. 814. 1942. 87 pp. 10 cents.

The bulletin treats of the problem of estimating the production and quality of the wheat crop according to crop-reporting districts of groups 10–20 counties within the wheat-belt States, in the light of the 1939 and 1940 field surveys.

The first section describes techniques employed in field and laboratory. Field work had been carried out by three (occasionally four) crews, each of two men in a car equipped with a crop meter. Assigned to work particular districts each crew selected its own itinerary so as to cover as nearly as practicable a grid-like pattern of highway, roughly equal to the length plus width of the counties individually. Frontage of wheat and total mileage

were recorded, and at intervals of ½ to 4 miles the nearest field was sampled by drawing therefrom two sampling units of 1/5000 acres each. An interesting comment is that the entire area clipped in Kansas in 1940 amounted to only 0.44 acre with an estimated standard error of mean yield per acre of less than 0.4 bushel.

Sample heads were sealed in prepared envelopes together with other information and mailed twice a day to the central laboratory for counting, measuring, threshing, and milling. The plan provided for timely "flash" reports on the production and quality. The average lapse of time in Kansas in 1940 between field sampling and release of information in mimeograph form was 6.4 days.

The second section deals with the efficiency of the sampling and costs for the information obtainable; and sources of bias. Analysis disclosed that the sampling interval and total mileage driven in each district should have been increased, while fewer fields need have been sampled. Another case study in sampling is thus added to the accumulating evidence that the aggregate of small samples, each of one or two sampling units from every subdivision of land area containing the crop, provides the estimate of greatest efficiency.

The authors find that it is more difficult to reduce systematic errors of observation to an acceptable minimum than the random errors of sampling.

The third and final section deals with forecasting the crop from sample measurements of head number and plant length taken 3–4 weeks before harvest. The method is not promising, however, because the regression equations differ from year to year.

The bulletin should prove very helpful in designing sample surveys of other annual crops and of perennial crops such as timber. Proofreading, however, seems not to have been taken very seriously. Rarely indeed does one find typographical errors in bulletins of the U. S. Department of Agriculture; yet this one contains several serious ones, as in the formula on page 18 and 23 and the italicized heading of the fourth column on page 25.

F. X. SCHUMACHER

Duke University

Some Factors Affecting the Supply of Milk and Milk Products in Nova Scotia, by Wilfred J. Garvin. Washington: The Catholic University of America Press. Studies in Economics, No. 4. 1941. viii, 155 pp.

This book "confines its attention to the dairy industry of Nova Scotia, the most populous of the Maritime Provinces . . . The interest of the present study centers around the fact that the total volume of milk production in Nova Scotia is substantially less than the total disappearance of milk and milk products within the Province. Consequently, it is chiefly concerned with the factors conditioning the volume of production, farm costs, and farm income."

Following an introductory chapter on the historical development and

present economic significance of the dairy industry in Nova Scotia, a brief review of some theoretical factors affecting milk supply and a description of the markets for different milk products is presented. Attempts are made to relate, by multiple regression analysis, butter production to various factors that might be expected to affect supply. Location of creameries is indicated and some idea of the extent of overlapping of the supply areas of these creameries is given. Estimates of costs of cream hauling and of plant operation are given.

The data on the costs of cream hauling and plant operation suffer, as do other parts of the book, from a lack of detailed factual information. Only summary figures are given and these appear to represent only the estimates of creamery managers as to their average unit costs, with no analysis of the items included in the total or the methods used to compute them.

On the basis of the analysis presented, it is concluded that "the solution of the present problem of low production, relative to local disappearance, is probably not to be advanced materially by such reductions in cost and changes in creamery organization as are practicably possible. Any economies that may be effected in this field would undoubtedly strengthen the competitive position of Nova Scotia farmers but they would probably be too small to constitute a substantial contribution to the solution of immediate problems." This statement as it stands may be true, but studies completed elsewhere have estimated the potential savings from route reorganization and adjustment of plant capacities to volume to be quite significant and worth securing.

Finally, after attention is paid to factors affecting hay prices and production costs, the following conclusions are reached: (1) that some opportunities exist for reducing the unit cost of creamery operations and for maintaining a grading system for cream purchased; (2) that unit costs of production should be reduced by (a) improving soil fertility, (b) enlarging cultivated acreages, (c) diversifying crops, and (d) utilizing land more intensively.

When the author refers to "the deficiency in the production, relative to provincial disappearance, of milk products in Nova Scotia" as a "problem" some may disagree. Indeed, it would appear that the real problems are those connected with farm organization and operation, and only in the event that changes in these would so modify total production as to change this "deficiency" into a surplus, would the relation of quantities produced to those consumed become a problem. In that case, Nova Scotia producers would be selling on a surplus rather than a deficit market and the present premiums they receive over central Canadian markets would be lost.

The study was presented as a doctoral dissertation. It brings together in one volume material on the dairy industry of Nova Scotia scattered in a number of different places. It presents an account of one phase of the economy of the Maritime Provinces of Canada.

Alan MacLeod

New England Research Council

Hospitals and Hospital Patients in New York City, by Neva R. Deardorff and Marta Fraenkel. New York: Welfare Council of New York City. Hospital Discharge Study, Volume One. 1942. xvii, 209 pp.

This is the first of a three volume analysis of hospital discharges in New York City during 1933 prepared by the Welfare Council of New York City with the assistance of state and city work relief authorities and the Work Projects Administration. Hospital records are kept primarily for administrative purposes and only those who have attempted to utilize such records in a study of illness can appreciate the almost herculean efforts required to abstract and summarize the records of 576,623 discharges.

After a brief discussion of the background and methods, material is presented concerning the hospitals cooperating in the study, the residence of patients and the amount of interborough and inter-health area "migration" for hospital service, the diagnosis of the disease or other condition for which treatment was given, certain demographic data, such as sex, age, color, and religion of the patients, the type of accommodation used and the pay status of the patients, the length of stay in the hospital and condition at time of discharge. These data are presented in 58 text and 12 appendix tables.

It is not possible to appraise fairly the results of this study on the basis of only the first of three contemplated volumes. However, a few general comments can appropriately be made at this time.

Certainly at present and probably for many years to come hospitalized illnesses cannot be accepted as representative of illnesses in the general population. Even in cities with the most complete hospital service probably not more than 10 to 15 per cent of all illnesses and 15 to 20 per cent of all disabling illnesses are hospitalized.

Nevertheless, current statistics of hospital admissions including such items as the number of admissions, the nature of the illness, the type of service obtained, and the financial status of the patients would be very useful in evaluating the adequacy of hospital service in spite of the fact that such statistics would not yield representative information concerning even the more serious illnesses.

The present study is confined to discharged patients. Consequently, it does not reveal the number of persons seeking hospital care during the year, that is the incidence of hospitalization, or the average number of patients in hospitals at a given time, the prevalence of hospitalization. Of course, over a sufficient period of time admissions and discharges must counterbalance but for short periods of time this is not necessarily true especially if hospitals for the chronically ill are included.

Perhaps the principal value of this study will be the experience gained in the attempt to utilize hospital records for other than strictly administrative purposes and the analysis of the possibility of initiating hospital morbidity reports. A discussion of these points to be published in Volume 3 will be eagerly awaited.

HAROLD F. DORN

United States Public Health Service

Income as an Index of the Fiscal Capacity of Michigan Counties, by Marvin A. Bacon. Ann Arbor: University of Michigan Press. Michigan Governmental Studies, No. 8. 1941. vii, 78 pp. 35 cents.

This is a welcome addition to the all-too-short list of studies published to date on measurement of fiscal capacity. The large and growing grants-in-aid for relief and other governmental functions make it increasingly important to determine a sound basis for distributing such aid. And the strong preference for equalizing rather than stimulating aids, as evidenced by recent developments in both state school and relief grants, makes it clear that need and fiscal capacity will be used more and more to determine the distribution of federal and state aids for education and welfare.

Measurements of need and fiscal capacity in actual use are far from perfect. There is neither general agreement as to the best tests nor adequate factual data on distribution of wealth and income either among states or within them. And there is danger that vested interests in a faulty formula will freeze an unsatisfactory distribution system unless progress in measuring fiscal capacity at least keeps pace with growing financial aid.

The author has limited himself to measuring the relative fiscal capacity of the counties of Michigan for the purpose of relief grants. The interest of the study lies, however, in the method, which has much wider application. State income has been broken down into agricultural, other private production, government, and all other. Each of these items has been been apportioned among the counties according to indirect measures of income—the amount of income itself not being available. The method varies for each class. Having estimated county income per capita, the author proceeds to estimate per capita "supernumerary income" by subtracting from per capita income a minimum-of-subsistence income corrected for differences in cost of living in the different counties.

Each step of these estimates has been fully described, and the many limitations of the original data, the method of estimating, and the validity of the checks are discussed. The author makes no exaggerated claims either for the method used or the specific results. In this respect the study is one of the best that the reviewer has seen.

The author concludes that supernumerary income per capita offers the best available index of fiscal capacity. He ranks per capita income excluding income from government second, per capita income from all sources third, assessed valuation per capita fourth, and total assessed valuation (the basis of state relief grants in use in Michigan) fifth. In making this selection the author has apparently overlooked one important factor in determining the fiscal capacity of counties, namely that these local governments must depend on a property tax, the base of which is more closely tied to assessed values than income—total or supernumerary. Fiscal capacity must depend on *taxable* income or wealth.

Mabel Newcomer

Vassar College

The Relation of Cost to Output for a Leather Belt Shop, by Joel Dean, With a Memorandum on *Certain Problems in the Empirical Study of Costs*, by C. Reinold Noyes. Technical Paper 2. New York: National Bureau of Economic Research. 1941. 74 pp. 50 cents.

This publication is certainly a combination of extremes. The first 52 pages contain the results of a thorough statistical study of the nature of cost curves in a belt shop and the last 20 pages contain a general critical examination of the theory of costs to see if it is sound enough to warrant an attempt to "test its validity" by quantitative analysis. The reviewer wishes that the two parts had been published separately. Even Noyes admits that his memorandum "is not entirely relevant to this particular study," and is really "an examination of the National Bureau's project and its method of attack, rather than one of this particular study." Noyes closes his memorandum with the statement:

> In these remarks it is far from my purpose to discourage efforts to establish by empirical studies truly scientific generalizations as to the relations of costs to output. . . . On the contrary, I consider this subject of costs to be perhaps the most potentially fruitful field that offers itself to real economic tillers of the soil. My only object is to call attention to the arduous job of clearing, draining, ploughing, harrowing, and sowing that must be done before it is going to be worth while to bring out the reaper.

In spite of this statement most readers are certain to reach the conclusion that Noyes thinks Dean used the reaper without waiting for economic theorists to do the "harrowing!"

The reviewer does not agree with Noyes that Dean should have waited until all of the principal economists agree upon a theory of cost. Noyes starts out by raising such questions as: What is the nature of a cost? Is it a concept? Is it a calculation? Or, is it a kind of event which occurs in economic life? However, the reviewer liked Noyes' statements that "economic theory has reneged" upon the issue of "the destination of costs which are not allocated," "the fact that these questions are treated in a certain way in accounting practice or in the theory of costs is not sufficient to settle the procedure for scientific purposes," and "when reducing operations one plant may retain men according to seniority while another retains the most efficient." The last statement has a bearing upon marginal cost.

Dean studied the effects of nine variables which "the management thought might affect cost." He concluded that he should use "a linear total combined cost function in the multiple regression analysis." The findings for combined cost, direct,[1] and overhead[2] cost are presented in both graphic and tabular form. The total, average, and marginal aspects of cost behavior are shown for the three independent variables which were significant. Limited space prevents detailed examination of the results and of the adjustments made

[1] Cement, direct labor, and leather.
[2] Fixed charges, indirect labor, repairs, and supplies.

for changes in wage rates and material prices. Original data were not presented in tabular form, so the accuracy of the work cannot be evaluated.[3] The study appears to be an excellent one. The reviewer would like to read Dean's statement of what changes (if any) he would make in his study if he could repeat it after reading Noyes' memorandum.

Harry Pelle Hartkemeier

University of Missouri

Operating Results of Department and Specialty Stores in 1941, by Charles A. Bliss, and *Department Store Financial Requirements 1920–1940*, by Pearson Hunt and Albert R. Koch, Bulletin No. 114. Boston: Harvard University Graduate School of Business Administration, Bureau of Business Research, 1942. vi, 62 pp. $2.50.

This year's edition of this series is one of the more important, because it serves as a bench mark by which to measure the influence on department store retailing of the General Maximum Price Regulation, of shortages of consumer merchandise, and of other war conditions. This is the twenty-second year the Harvard Business School has published this analysis—which from the start has been a cooperative venture with the National Retail Dry Goods Association. This series, then, offers one of the most authentic and most historically complete series of retailing data available. The cooperation of retail firms—figures are based on reports from 492 companies operating 651 stores, and representing 37.5 per cent of the total volume of sales of department stores and departmentized specialty stores in the United States—and the year-by-year refinements in method of analysis make comments on the adequacy and on the method of treatment superfluous.

A note of caution in the text of the report reminds readers that the data in the report have been submitted, typically, by larger stores, and by stores whose executives consider store operation from the point of view of control. Although the data may not be typical of department store retailing as a whole, it is the feeling of the reviewer that the report reflects operating conditions of the most significant and most influential stores.

It is difficult to summarize summary tables. Several points of general interest should be stressed, however. In 1941, department store and specialty store gross margin was substantially higher than in any of the other twelve comparative years covered in the report; and for department stores, total expense was lower than for any year since 1930, and net profit was substantially higher than for any of the other years shown. However, a significant and sinister increase in the total expense per transaction appeared; only increased size of transaction and increased numbers of transactions made the profit showing possible. If department stores suffer a decrease in the average amount of sale, or in the number of sales, it seems likely that they will lose their recent gains.

[3] In Table 9 (p. 34) 0.28 should be changed to .028. If the material on page 41 does not contain a misprint it needs clarification.

Financial Requirements of Department Stores, 1920–1940 (added to the report) is a cooperative study with the National Bureau of Economic Research. This report analyzes the financial statements of fourteen large department stores over a twenty year period to ascertain the sources of funds for store operation. Figures used for this analysis are appended to the report. Conclusions are that the growth of consumer instalment credit has pushed a growing financing burden on department stores; that inventories are being held increasingly to close relationship with sales, that an increasing proportion of reserves is being held in cash, but that the proportion of working capital to total assets has decreased; and that fixed assets have increased over the twenty-year period both absolutely and in relation to total assets.

LAWRENCE C. LOCKLEY

The Curtis Publishing Company

French Predecessors of Malthus, by Joseph H. Spengler. Durham: Duke University Press. 1942. ix, 398 pp. $4.50.

As the author states in the preface, there has been a need for a detailed and authoritative study of French population and wage theory of the pre-nineteenth century period, especially since many of the original materials are not readily available and the secondary sources are generally too specialized or too summary in character. This need is well filled by the present volume which gives detailed summaries of both major and minor French contributions to early population literature, being in effect a fuller documentation of the author's earlier work, *France Faces Depopulation*.

The scope of coverage is greater than indicated by first glance at the title, for population and wage theory are broadly defined and the writers dealt with are the predecessors of Malthus in point of time, not his precursors in point of thought. To judge from its form of organization the book is designed for reference use, the treatment of the subject matter being neither strictly chronological nor in terms of theoretical sequences. The various writers are dealt with in turn, full abstracts of their works being given, but further organization is secured through their grouping into several schools of thought, numerous cross-references, and a detailed index.

It is shown that although the subject of population received considerable attention from French writers of the seventeenth and eighteenth centuries, few anticipated in any degree the Malthusian doctrine. Opinion in the earlier period is seen to have been quite unanimously in favor of population growth as the sure means to national strength and prosperity. During the eighteenth century the Physiocrats and other economists came to emphasize the importance of resources and industry as sources of wealth, at the same time that the growing liberal and social point of view stressed considerations of individual welfare. Elements of the Malthusian doctrine are shown to have appeared here and there but only a few writers, notably Brückner and Mann, were in a real sense precursors of Malthus. Although foreigners such as

Arthur Young could conclude that France was suffering from overpopulation, the existing social ills were currently attributed to faulty government, defective economic organization, and maldistribution of wealth.

A unified theory of wages was not to be developed until later. In the discussion of the subject in eighteenth century France, quite as in contemporary England, there are shown to have been elements of supply and demand or subsistence, wage-fund, and productivity theories.

E. P. Hutchinson

Washington, D. C.

On the Principles of Statistical Inference, by Abraham Wald. Notre Dame, Indiana: Notre Dame Mathematical Lectures, No. 1, 1942. 47 pp.

Two problems of statistical inference have received considerable attention during the last twenty or thirty years: the testing of statistical hypotheses and the problem of estimation. In this booklet, containing four lectures delivered by Dr. Abraham Wald at Notre Dame University in February, 1941, these two problems are stated, as well as the general problem of statistical inference of which they are special cases; a brief general review of the solutions obtained for them is given; and an approach to the general problem, following concepts developed by Dr. Wald, is indicated. For the mathematician or statistician who wishes to get a brief over-all review of the progress that has been made in mathematical statistics, and of some of the major problems still unsolved, Dr. Wald's paper may be highly recommended.

The contents are divided into six chapters: the first is an introduction which defines the terminology used, presents the general problem of inference, and gives the special cases of estimation and of testing a statistical hypothesis. The second chapter considers the concepts introduced by J. Neyman and E. S. Pearson in developing their theory of testing an hypothesis regarding the value of a single unknown parameter, and the third chapter gives R. A. Fisher's approach to unique or point estimation of a parameter. Estimation of the value of a parameter by a confidence interval, following Dr. Neyman's theory, is treated in Chapter IV. Unfortunately, the briefness of the publication has led to the omission of many interesting topics, notably testing composite hypotheses, estimating one unknown parameter when there are several unknown parameters involved in the probability distribution, and Fisher's theory of fiducial limits. However, despite their briefness and the consequent omissions, these four chapters give a very good outline of the basic ideas used in the theories of Fisher, Neyman and Pearson. Dr. Wald's own recent work in developing asymptotically most powerful tests and asymptotically shortest confidence intervals is treated in Chapter V, and several interesting results are given. In the last chapter, the author considers the general problem of statistical inference, outlines the steps that must be taken to solve it, and gives his approach to the solution, through the consideration of the risk of loss involved in making an incorrect inference.

Although briefness has led to omissions of subjects and of detailed develop-

ments, this very briefness has the advantage of letting the reader obtain quickly an excellent grasp of the entire field of statistical inference, and of the concepts used in modern statistical theory.

MARK EUDEY

Statistical Laboratory,
University of California,
Berkeley

Changes in Distribution of Manufacturing Wage Earners, 1899–1939. Washington: Bureau of the Census and Bureau of Agricultural Economics, 1942, 268 pp. 30¢.

This report, prepared by Harold D. Kube and Ralph H. Danhof, is a study of the geographic distribution of manufacturing wage earners in the United States and results from the cooperative study that the Bureau of the Census and the Bureau of Agricultural Economics started early in 1940.

The purpose of the study was to develop accurate manufacturing wage-earner data to form the basis for further analytical work. No part of the report is designed to explain the reasons for changes in the distribution of manufacturing wage earners or the economic consequences of such changes. The study is based primarily upon the Censuses of Manufactures and the maps, charts, and tables forming the body of the report are presented in three sections.

Section I describes changes in the distribution of manufacturing wage earners between 1899 and 1939 for the Nation as a whole, the 9 geographic divisions, the 48 states, the 33 industrial areas, and various city groups. The importance of the changes in the distribution of wage earners is indicated by relating them wherever possible with corresponding changes in the population of each area. Sufficient data are found in this section to prepare analytical studies which would trace the growth process of the present manufacturing wage-earner employment in each area.

Section II contains maps, charts, and tables showing changes in the distribution of wage earners between 1929 and 1939 in geographic divisions, industrial areas, and city groups for 24 manufacturing industries considered important to national defense. Wage-earner data are indicated on both the 1929 and the 1939 maps by counties and the relative importance of each county is designated by shading or ruling for 5 class groups. The limits of the categories were set wide enough that information could be given for all counties in which the industry occurs without violating the disclosure rule which forbids the Bureau of the Census to reveal data of a given establishment. It is possible, however, to discern readily whether the industry is of minor, small, medium, considerable, or great importance in terms of number of wage earners in any part of the country. Charts and tables are presented to support the maps by showing the distribution of wage earners by geographic divisions, industrial areas, and 5 groupings of cities for each manufacturing census between 1929 and 1939 inclusive. Such information should contribute much to any regional study of the selected industries.

Section III contains source material in two sets of tables. The first set shows the number of manufacturing wage earners employed between 1899 and 1939 for the United States, geographic divisions, states, counties, industrial areas, and groups of cities. The second set comprises lists of the cities included in the various city groups based on the 1930 and 1900 census of population. The source data on the first set have been adjusted so that they are as nearly as possible comparable over the 40-year period. The authors hope that these statistics will be the basis for future historical analyses of the development of the geographic structure of American manufacturing activity.

The county and city wage-earner data published in this report should be of particular usefulness to statisticians, economists, and other students of manufacturing data because they are probably the only set of data available in which previously published census data were revised to make the county and city statistics more comparable over the 1899–1939 period. Many changes were made in the earlier published data and, in addition, many new figures were added by estimating the number of wage earners for those counties and cities for which separate information was not previously reported.

Changes in scope of the census, definition of average number of wage earners, areas included in city or county boundaries, and coverage of the several censuses were considered in correcting the wage-earner data for comparability. The estimates of county and city statistics to overcome the lack of completeness in published census data were made according to accepted statistical procedures. Good use seemed to have been made of the information available and it is unlikely that the error in these estimates will affect the usefulness of the data. The authors indicate that the possible margin of error is usually rather small because the upper and lower limits of their estimates could be determined with reasonable accuracy.

The report is highly recommended as a source book of statistics to all persons interested in analyzing changes in the distribution of manufacturing wage earners or in making regional studies of selected manufacturing industries. Its small cost for the wealth of available information makes the bulletin especially attractive.

Oscar L. Endler

National Resources Planning Board

Industrial Statistics, by H. A. Freeman, New York: John Wylie & Sons, Inc., 1942. ix, 178 pp. $2.50.

Production could be increased by thirty per cent, and costs could be decreased by a like amount, with a corresponding improvement in quality, if the principles of quality control were to be applied to production, in place of the conventionally accepted inspection methods now used. By quality control is implied establishing a qualia level and so regulating and controlling the tools of production and the manufacturing process so that this qualia is always obtained.

At present the accepted procedure in most industries is to make the product and then submit it to an inspector for acceptance or rejection. This guarantees the consumer a quality product but does not insure the producer against loss from spoilage and defective workmanship. The cost of this spoilage is ultimately absorbed by the consumer which results in higher prices. Production is also restricted by the amount of spoilage since this requires reworking of the rejected product or scrapping of the defective units and reordering new materials in order to make the units over again. This waste of man power and machine hours is irreplaceable.

It has long been recognized that conventional inspection methods, while satisfactory to the consumer, did not contribute to increased production or lower costs for the producer. The amazing thing is that manufacturers have been satisfied to organize and hire a production force and then parallel this organization with an inspection department with authority to reject the product made by the manufacturing organization. It is surprising that the principles of *preventive* inspection have not received more attention in view of the tremendous losses due to spoilage during manufacturing.

Dr. Shewhart, of the Bell Laboratories, recognized this problem a few years ago, and as a result of his pioneering efforts, a considerable amount of work has been done in the last half dozen years to replace the post mortem inspection operation with proper controls during manufacturing so that the desired quality level might be maintained during the manufacturing cycle, thus reducing rejections to a minimum.

In his new book, Professor Freeman has attacked the problem of quality control in a novel but effective manner. He starts out by determining the answer to the question: Of two or more lots of materials, did they all come from the same universe, or, in the language of the shop man, were they all produced under identical conditions? In other words, Professor Freeman gives a mathematical solution for the age-old problem of probability: Is the difference between lots what might be expected due to chance, or are the lots significantly different from each other?

From this test, Professor Freeman proceeds to the use of correlation and the analysis of variance to be applied to manufacturing operations during the manufacturing cycle. Unfortunately, on this subject, Professor Freeman does not have too much to say about the control of quality applied to the metal working industries. However, the quality control engineer will find this study a practical guide, though somewhat too mathematical for the ordinary shop man. Professor Freeman attempts to get around this difficulty by segregating into separate sections the more theoretical treatment of the mathematics of the problem.

The tables at the end of the book are a boon to the serious student of this subject, since in the past he has had to go to have a dozen books to get this information.

JOSEPH MANUELE

Westinghouse Electric and Manufacturing Company

PUBLICATIONS RECEIVED*

Anthony, Donald, et al. *How Collective Bargaining Works.* New York: The Twentieth Century Fund. 1942. xxviii, 986 pp. $4.00.

Bagley, William Chandler, Jr. *Soil Exhaustion and the Civil War.* Washington: American Council on Public Affairs. 1942. xi, 101 pp. $2.00.

Beebe, Gilbert Wheeler. *Contraception and Fertility in the Southern Appalachians.* Baltimore: The Williams & Wilkins Company. 1942. vii, 274 pp. $2.50.

Brown, Earl, and Leighton, George R. *The Negro and the War.* New York: Public Affairs Committee. 1942. 32 pp. 10 cents.

Crum, William Leonard, Fennelly, John F., and Lawrence, Howard S. *Fiscal Planning for Total War.* New York: National Bureau of Economic Research. 1942.

Deardorff, Neva R., and Fraenkel, Marta. *Hospitals and Hospital Patients in New York City.* New York: Welfare Council of New York City. Hospital Discharge Study, Volume One. 1942. xvii, 209 pp.

East, W. Norwood. *The Adolescent Criminal.* London: J. & A. Churchill, Ltd. 1942. xi, 327 pp. 45 shillings.

Faville, David E., Jones, Dix M., and Sonne, Richard B. *Merchandise Availability in Utah.* Stanford University: Stanford University Press. 1942. xvi, 229 pp. $3.00.

Field, Kenneth, Editor. *American Finance.* New York: American Finance Association. Abstracts of Papers Presented at the 1941 Annual Meeting of the Association. 1941. v, 57 pp.

Holmes, Maurice C. *Outline of Probability and Its Uses.* Minneapolis: Burgess Publishing Company. 1936. viii, 119 pp. $1.50.

Hoover, Edgar M., and Dean, Joel, Editors. *Readings in the Social Control of Industry.* Philadelphia: The Blakiston Company. 1942. viii, 494 pp. $2.75.

Jacoby, Neil H., and Saulnier, Raymond J. *Term Lending to Business.* New York: National Bureau of Economic Research. 1942. xvii, 163 pp. $2.00.

Marlio, Louis. *Will Electric Power Be a Bottleneck?* Washington: The Brookings Institution. 1942. 67 pp. 50 cents.

Molina, E. C. *Poisson's Exponential Binomial Limit.* New York: D. Van Nostrand Company, Inc. 1942. 47 pp. $2.75.

Palmer, Gladys L., assisted by Cohn, Samuel M. *War Labor Supply Problems of Philadelphia and Its Environs.* Philadelphia: University of Pennsylvania, Industrial Research Department of Wharton School of Finance and Commerce. Research Report No. 5. 1942. vii, 54 pp. 50 cents.

Pollock, James K., and Eldersveld, Samuel J. *Michigan Politics in Transition, An Areal Study of Voting Trends in the Last Decade.* Ann Arbor: University of Michigan Press. 1942. 74 pp.

Stigler, George J. *The Theory of Competitive Price.* New York: The Macmillan Company. 1942. vii, 197 pp. $3.00.

Toner, Jerome L. *The Closed Shop.* Washington: American Council on Public Affairs. 1942. viii, 205 pp. $3.25.

Treloar, Alan E. *Random Sampling Distributions.* Minneapolis: Burgess

* Annual reports and publications presenting statistics collected at regular intervals have been omitted from this list. Some items of minor interest to statisticians have also been omitted. The contents of periodical publications are not listed, but the attention of the reader is directed to the lists of articles in current publications which are to be found in the *Revue de l'Institut Internationale de Statistique, Journal of the Royal Statistical Society, American Economic Review, Population Index, Transactions of the Actuarial Society of America, The Record of the American Institute of Actuaries,* and *Sankhyā, The Indian Journal of Statistics.*—Editor.

Publishing Company. 1942. 94 pp. $2.25.

Yoder, Dale. *Personnel Management and Industiral Relations.* New York: Prentice-Hall, Inc. 1942. xxii, 848 pp. $5.65.

Government Publications

U. S. Department of Commerce, Bureau of the Census, and Department of Agriculture, Bureau of Agricultural Economics. *Changes in Distribution of Manufacturing Wage Earners 1899–1939,* prepared by Harold D. Kube and Ralph H. Danhof. 1942. vii, 268 pp. 30 cents.

Minnesota Resources Commission. *Minnesota Incomes 1938–39.* A Report of the Minnesota Income Study. Saint Paul: Minnesota Resources Commission. Vol. I. 1942. 177 pp.

——— Vol. II. 1942. 369 pp.

——— Vol. III. 1942. 329 pp.

——— Vol. IV. 1942. 56 pp.

Social Security Board, Federal Security Agency, Bureau of Research and Statistics. *Statistics of Family Composition in Selected Areas of the United States, 1934–36.* Volume 10. Los Angeles, California. 1942. xxviii, 303 pp.

The Annals of Mathematical Statistics

THE OFFICIAL JOURNAL OF THE INSTITUTE
OF MATHEMATICAL STATISTICS

Edited by S. S. WILKS, A. T. CRAIG, and J. NEYMAN

CONTENTS OF VOLUME XIII, NUMBER 3, SEPTEMBER 1942

	Page
Additive Partition Functions and a Class of Statistical Hypotheses. J. Wolfowitz	247
On the Theory of Testing Composite Hypothesis with One Constraint. Henry Scheffé	280
On the Problem of Multiple Matching. I. L. Battin	294
On the Choice of the Number of Intervals in the Application of the Chi Square Test. H. B. Mann and A. Wald	306
Limited Type of Primary Probability Distribution Applied to Annual Flood Flows. Bradford F. Kimball	318
Linear Restrictions on Chi Square. Franklin E. Satterthwaite	326
Systems of Linear Equations with Coefficients Subject to Error. A. T. Lonseth	332
On Mutually Favorable Events. Kai-Lai Chung	338
Observations on Analysis of Variance Theory. Hilda Geiringer	350

Subscription Rate: $5.00 per annum

Address orders for subscriptions and back numbers to Professor Edwin G. Olds, Secretary, Institute of Mathematical Statistics, Carnegie Institute of Technology, Pittsburgh, Pennsylvania

Copies of the Cumulative Index of Volumes I-X of the *Annals of Mathematical Statistics* can be obtained at $.50 each from the Secretary.

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

OFFICERS AND EDITORS FOR THE YEAR 1942

President

ALFRED J. LOTKA
Metropolitan Life Insurance Company
New York City

Directors

W. LEONARD CRUM
Harvard University
Cambridge, Massachusetts

LOWELL J. REED
The Johns Hopkins University
Baltimore, Maryland

WALTER A. SHEWHART
Bell Telephone Laboratories
New York City

F. LESLIE HAYFORD
General Motors Building
New York City

WINFIELD W. RIEFLER
The Institute for Advanced Study
Princeton, New Jersey

FREDERICK F. STEPHAN
Cornell University
Ithaca, New York

Vice-Presidents

CALVERT L. DEDRICK
Bureau of the Census, Washington, D. C.

Collection and Classification of Data and Administration of Statistical Agencies

NORMAN J. SILBERLING
Stanford University, California

Facts and Methods Bearing upon Economics and Economic Theory

WALTER BARTKY
University of Chicago, Chicago, Illinois

Statistical and Actuarial Methods and Technique, and the Teaching of Statistics

HENRY B. ARTHUR
Swift & Company, Chicago, Illinois

Facts and Methods Pertaining Primarily to Business

HOWARD B. MYERS
War Production Board
Washington, D. C.

Facts and Methods Pertaining to Sociology, Social Welfare Problems, and Labor Statistics

HAROLD V. ROELSE
Federal Reserve Bank of New York
New York City

Facts and Methods Pertaining to Financial Institutions

FRANK W. NOTESTEIN
Office of Population Research, Princeton University
Princeton, New Jersey

Facts and Methods Related to Biometry, Vital Statistics, Psychology, and Education

ALFRED N. WATSON
Curtis Publishing Company
Philadelphia, Pennsylvania

Facts and Methods Pertaining to Marketing

Secretary-Treasurer

R. L. FUNKHOUSER*
The American University
Washington, D. C.

Editor

FRANK A. ROSS
The American University
Washington, D. C.

Review Editor

GLENN E. MCLAUGHLIN
The American University
Washington, D. C.

Associate Editors

F. LESLIE HAYFORD
General Motors Building
New York City

LOWELL J. REED
The Johns Hopkins University
Baltimore, Maryland

GLENN E. MCLAUGHLIN
The American University
Washington, D. C.

FREDERICK F. STEPHAN
Cornell University
Ithaca, New York

Biometrics Section

Chairman

C. I. BLISS
The Connecticut Agricultural Experiment Station
New Haven, Connecticut

Secretary

JOSEPH BERKSON
Mayo Clinic
Rochester, Minnesota

* On leave of absence.

CHAPTERS AND DISTRICT REPRESENTATIVES

ALBANY CHAPTER—Edward T. Frankel, *491 State Street, Albany, New York*

ASOCIACIÓN CUBANA DE ESTADÍSTICA CHAPTER—Jacob Brandon, *P. O. Box 2187, Havana, Cuba*

ATLANTA CHAPTER—Paul G. Barnett, *Southern Bell Telephone and Telegraph Company, 1501 Hurt Building, Atlanta, Georgia*

AUSTIN CHAPTER—John R. Stockton, *University of Texas, Austin, Texas*

BOSTON CHAPTER—Roswell F. Phelps, *Massachusetts Department of Labor and Industries, Boston, Massachusetts*

CENTRAL INDIANA CHAPTER—Clausin D. Hadley, *Indiana University, Bloomington, Indiana*

CENTRAL NEW JERSEY CHAPTER—Clarence D. Long, Jr., *Institute for Advanced Study, Princeton, New Jersey*

CHICAGO CHAPTER—John F. Reinboth, *Commonwealth Edison Company, 72 West Adams Street, Room 1548, Chicago, Illinois*

CINCINNATI CHAPTER—F. E. Wolfe, *Procter and Gamble Company, Cincinnati, Ohio*

CLEVELAND CHAPTER—Edward A. Stephen, *The Ohio Bell Telephone Company, Cleveland, Ohio*

COLUMBUS CHAPTER—Paul J. Daugherty, *Ohio Chamber of Commerce, 1325 Huntington Bank Building, Columbus, Ohio*

CONNECTICUT CHAPTER—Louis M. Nichols, *General Electric Supply Corporation, 1260 Boston Avenue, Bridgeport, Connecticut*

DETROIT DISTRICT CHAPTER—Andrew T. Court, *1052 Seminole Avenue, Detroit, Michigan*

HARRISBURG STATISTICAL SOCIETY—Henry W. Van Pelt, *Pennsylvania Department of Internal Affairs, Harrisburg, Pennsylvania*

MADISON WISCONSIN CHAPTER—Churchill Eisenhart, *University of Wisconsin, Madison, Wisconsin*

MONTGOMERY CHAPTER—Leonard V. Phelps, *Alabama Department of Health, Montgomery, Alabama*

NEW YORK DISTRICT CHAPTER—Mrs. Helen Slade, *400 East 57th Street, New York City*

NORTH CAROLINA CHAPTER—Richard L. Anderson, *North Carolina State College, Raleigh, North Carolina*

PHILADELPHIA CHAPTER—J. Frank Rehfuss, *Department of Research and Statistics, Federal Reserve Bank of Philadelphia, Philadelphia, Pennsylvania*

PITTSBURGH CHAPTER

SAN FRANCISCO CHAPTER—William A. Sturm, *California State Chamber of Commerce, 350 Bush Street, San Francisco, California*

UNIVERSITY OF ILLINOIS CHAPTER—Garret L. Jordan, *University of Illinois, Urbana, Illinois*

WASHINGTON STATISTICAL SOCIETY—Harold D. Kube, *War Production Board, New Social Security Building, Washington, D. C.*

THE ANNALS **November, 1942**

LABOR RELATIONS AND THE WAR

Edited by HERMAN FELDMAN, Ph.D., Professor of Industrial Relations, Dartmouth College, Hanover, New Hampshire

¶The outline of the volume calls for consideration of the subject under the following heads: Recent Developments in Labor Organization; Issues of Wartime Labor Standards; Labor's Cooperation in the National Effort; Industrial Peace and National Labor Policy; Planning Against Postwar Labor Tension. A selected bibliography is included.

$2.00

($1.00 each for additional copies to members of the Academy)

THE AMERICAN ACADEMY OF POLITICAL AND SOCIAL SCIENCE
3457 Walnut Street, Philadelphia, Pa.

JOURNAL OF FARM ECONOMICS

Published by THE AMERICAN FARM ECONOMIC ASSOCIATION

Editor: H. B. PRICE

University of Kentucky Lexington, Kentucky

Volume XXIV, November 1942, Number 4

Some of the feature articles are:

Bases for Controlling Agricultural Prices ..... GEOFFREY SHEPHERD

Agricultural Income and The Export Market, 1910-1940 ..... MARION CLAWSON AND JOHN D. BLACK

Wartime Wheat Policy in Canada .. H. S. PATTON

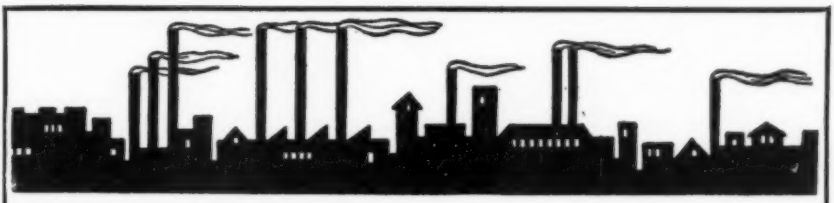
Cross-Sectional and Case-Grouping Procedures in Research Analysis.. LEONARD A. SALTER, JR.

Hired Farm Labor in World War II ..... HARRY SCHWARTZ

Emergency Control in the Farm Real Estate Market ..... M. M. REGAN AND FRED A. CLARENBACH

This Journal, a quarterly, contains in addition, notes, reviews of books and articles, and a list of recent publications. It is published in February, May, August, and November by the American Farm Economic Association. Yearly subscription $5.00.

Secretary-Treasurer: ASHER HOBSON
Department of Agricultural Economics
University of Wisconsin, Madison, Wisconsin



Statistical Techniques for Controlling War Production Business Operations

PRACTICAL BUSINESS STATISTICS

By Frederick E. Croxton and Dudley J. Cowden

Designed as Basic Text for College Courses in BUSINESS STATISTICS

ADOPTED BY

- Yale University
- University of Chicago
- Duke University
- University of Detroit
- DePaul University
- Detroit Inst. of Technology
- Rutgers University
- Texas A & M College

and many others!

With the speed-up in the production of war goods, modern statistical techniques for controlling business operations have assumed increasing importance. Today, more than ever, the industrial executive must have a graphic presentation of the facts and figures of his business to assist him in efficient management.

PRACTICAL BUSINESS STATISTICS explains clearly and simply the techniques for collecting, presenting, and analyzing business data. It shows how to use statistical procedures to control manufacturing costs, quality and quantity of materials, labor turnover, budgets, purchases, sales, profits, and other factors that enter into wartime business operations. Over 300 charts, graphs, tables, and forms are introduced to make everything crystal clear for the student.

The book widely adopted by leading colleges and universities, the names of some of which are listed in the column at the left. Try it in your own classes.

520 Pages • College List $3.50

Send for Your Examination Copy Today

PRENTICE-HALL, INC. ★ NEW YORK

Outstanding McGraw-Hill Books

FUNDAMENTAL STATISTICS IN PSYCHOLOGY AND EDUCATION

By J. P. GUILFORD, University of Southern California. 333 pages, 6 x 9. $3.25

Covers routine procedures as well as more recently developed ones, and underscores interpretation of statistical results sufficiently to serve as a handbook for the investigator in social or biological sciences. The treatment presupposes no previous knowledge of statistics.

STATISTICAL PROCEDURES AND THEIR MATHEMATICAL BASES

By CHARLES C. PETERS and WALTER R. VANVOORHIS, Pennsylvania State College. 516 pages, 6 x 9. $4.50

Designed to bridge the gap between elementary courses in which the formulas are given from a purely authoritative viewpoint, and the original monographic contributions which are so often highly mathematical in character. The authors bring together and synthesize the classical statistics and the new developments worked out by R. A. Fisher.

ELEMENTARY SOCIAL STATISTICS

By THOMAS C. MCCORMICK, University of Wisconsin. 353 pages, 6 x 9. $3.00

Aims to give professional sociologists an adequate background for later training in quantitative research methods. Part I considers the problems that are encountered when the quantitative method is applied to sociological data. Part II selects those elementary statistical methods that are most useful to the sociological investigator.

BUSINESS STATISTICS

By JOHN R. RIGGLEMAN, Consulting Economist and Statistician, and IRA N. FRISBEE, University of California at Los Angeles. Second edition. 790 pages, 6 x 9. $4.00

This leading text focuses the study upon the type of work that the practicing statistician is called upon to do, such as business forecasting, budgeting, production control, marketing analysis, etc.

STUDENT'S MANUAL FOR BUSINESS STATISTICS

By IRA N. FRISBEE and WILLARD N. ANDERSON, University of California at Los Angeles. 250 pages, 8½ x 11. $2.00

Send for copies on approval

McGRAW-HILL BOOK COMPANY, INC.

330 West 42nd Street **New York, N.Y.**

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

THE ANALYSIS OF ECONOMIC TIME SERIES

By PROF. HAROLD T. DAVIS, Northwestern University

This volume first reviews the historical development of the problem. Techniques of harmonic analysis, in particular, and of regression analysis, in general, are developed. Numerous periodograms are given together with an analysis of the conclusions to be derived from them. Problems associated with trends, forecasting, random statistical series, and cycle analysis are treated. Applications are made to the equation of exchange, stock prices, the distribution of income, interest rates, the building cycle, and other economic Phenomena. (Monograph Number 6 of the Cowles Commission for Research in Economics.)

This book contains 620 pages
Size 6½ x 9½. Price $5.00

THE THEORY OF ECONOMETRICS

By PROF. HAROLD T. DAVIS, Northwestern University

This book, a mathematical and statistical treatment of economics, is written in the form of a textbook. It is divided into two parts, one devoted to Economic Statics, and the other to Economic Dynamics.

This book contains 482 pages
Size 7 x 10. Price $4.00

THE PRINCIPIA PRESS, INC. Bloomington, Ind.

AMERICAN MARKETING ASSOCIATION

An Association for the Advancement of Science in Marketing

THE American Marketing Association, formed by merging the American Marketing Society with the National Association of Marketing Teachers, began its existence on January 1, 1937, and now has over 1000 active members in all parts of the country.

The principal purposes of the Association are to foster scientific study and research in the field of marketing, to improve the methods and technique of market research and to develop better public understanding and appreciation of marketing problems.

Active Chapters have been established and hold regular meetings in New York, Philadelphia, Washington, Chicago, San Francisco and Los Angeles, Local Chapters are being organized in other cities.

The official publication of the Association, *The Journal of Marketing*, is issued quarterly. Now in its sixth year, the Journal is regarded by marketing practitioners as a source of authoritative marketing information. Many of the contributing authors are leaders in their field of endeavor.

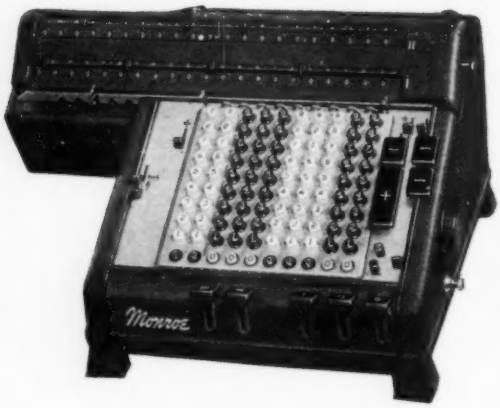
The full membership rate is $5 annually. This includes a subscription to *The Journal of Marketing*. For further particulars, please address the Secretary.

President	*Treasurer*	*Secretary*
Dr. VERGIL D. REED	HOWARD W. GREEN	Prof. ALBERT HARING
4131 Census Building	1001 Huron Road	Indiana University
Washington, D.C.	Cleveland, Ohio	Bloomington, Ind.

Monroe Adding-Calculator

STATISTICAL MODEL AA-1

DESIGNED BY statisticians based upon over a quarter of a century of service and experience in the use of Monroe Adding Calculators



DESIGNED FOR statisticians in the educational governmental research and business statistics fields

DURING the war this statistical Monroe Adding-Calculator, as well as other models can only be supplied to the fighting forces and essential war industries. In the interim Monroe is prepared to render maintenance service to lengthen the life of your equipment and keep it in condition to handle the following types of statistical calculations.

- Individual X and Y for proof.
- Summations X, Y, X^2, Y^2, XY.
- Single keyboard for all Set-ups.
- Exclusive squaring lock providing single set-up for X and Y.
- Automatic carriage positioning for X and Y.
- Automatic Negative Multiplication.
- Automatic short-cut for X and Y.
- Exclusive increased dial capacity provides for unlimited pairs of observations.
- Extreme flexibility for all formulae including computation for value of r.

MONROE CALCULATING MACHINE COMPANY, INC., Orange, N. J.

Consult nearest Monroe office or write to
Educational Department, Orange, N. J.

Important New Books for the Library

VON MERING—

The Shifting and Incidence of Taxation

This new book meets a long existing need of practical tax analysts and students of public finance for a thoroughly modern treatment of the principles of tax shifting. The analysis is up-to-date in economic methodology and theory and includes the most recent contributions to the literature of the subject. The author has contributed in addition, some novel ideas and new approaches to the theory of tax shifting which economists generally should find of much interest. The book achieves comprehensive coverage of different types of taxes by analytical approach rather than by classification of taxes according to their superficial resemblances. The marginal analysis proceeds from rigorously defined conditions of competition, monopoly and monopolistic competition, and the limitations of the theoretical results in the light of complex conditions and important "non-economic" considerations are carefully stated. Agreeable style of exposition relieves the necessary rigor of the treatment. The literary and diagrammatical presentation is entirely independent of the supplemental algebraic demonstrations which are introduced in small type.

By OTTO VON MERING, *Tufts College*

264 Pages, $3.25 (1942)

HALM—

Monetary Theory (A Modern Treatment of the Essentials of Money and Banking)

This book is a realistic theoretical treatment of all the major problems of monetary analysis and policy. The monetary analysis is integrated with the "real" or "non-monetary" aspects of economic relations and rises far above the confusions and shortcomings of one-sided monetary interpretations. The chapters on the loanable funds theory of interest; the interrelations of money, credit, interest and capital; money and the business cycle, and the problem of saving and investment are of special interest and highly useful. A work of thorough scholarship, its compact organization and simplicity of style is a conspicuous merit of the book. Practical bankers and financial administrators as well as students of economics will find this volume up-to-date and a most helpful aid to their understanding of monetary problems.

By GEORGE N. HALM, *Tufts College*

347 Pages, $3.50 (1942)

THE BLAKISTON COMPANY, Philadelphia

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*